Pattern Recognition for Weather Phenomena in Climate Data

Thesis submitted in accordance with the requirements of
the University of Liverpool for the degree of Doctor in Philosophy by

Grzegorz Muszynski

February 2021

"Each piece, or part, of the whole of nature is always merely an approximation to the complete truth, or the complete truth so far as we know it. In fact, everything we know is only some kind of approximation because we know that we do not know all the laws as yet."

R. P. Feynman

Acknowledgements

The path to this degree has started long before my PhD studies. I have been fortunate to meet and to have the support of many amazing people in this journey.

My first thanks go to my supervisors Dr Vitaliy Kurlin and Dr Michael Wehner who gave me a lot of valuable advice and freedom while carrying out my PhD research. I also thank Dr Karthik Kashinath and Dr Prabhat Ram for their mentorship during my stays at LBNL and for introducing me to the exciting world of climate science, atmospheric science, and supercomputing. I am grateful to them for their support in writing this thesis, papers, and their guidance.

I am also grateful to my PhD viva examiners: Professor Gabriele Hegerl and Professor Frans Coenen for evaluating my research and their valuable comments on the thesis. Furthermore, I appreciate great support of Professor Paul Spirakis and Professor Andrew Morse during my PhD studies.

I am thankful for being able to collaborate with Dr Jan Balewski, Dr Dmitriy Morozov, and Burlen Loring. Thank you, Jan and Burlen, for devoting so much time to me over the last few years. I have certainly learned from you a lot.

I want to thank all my friends and colleagues at Liverpool, NERSC, LBNL, UC Berkeley for their friendliness and helpfulness all these years, in random order: Dr Arturo Fernandez, Mahesh Balasubramanian, Bharath Ananth, Dr Mirosław Salamończyk, Dr Adam Rupe, Dr Chiyu (Max) Jiang, Yue (Luna) Huang, Professor Floriana Petrone, Haris Moazam Sheikh, Dr Adrian Albert, Dr Ana Catarina Silva, Fahim Chowdhury, Dr Jinlong Wu, Evan Racah, Dr Paweł Gniewek, Krasen Samardzhiev, Dr Gregory Palmer, Professor Gleifer Vaz Alvez, and Dr Steven Farrell.

I also want to thank all the people I met in these years of travelling to conferences, summer schools, and seminars: Dr José Luis Licón-Saláiz, Dr Barbara Giunti, Dr Henri Riihimäki, Dr Stephan Zhechev, Dr Kristóf Huszár, Dr Marie McGraw, Dr Ben Toms, and Professor Henry Adams for the helpful math and climate science discussions.

This PhD research would not have been possible without the financial support of NERSC and Intel. I would like to thank the funding partners — the NERSC-Intel Big Data Center and the Intel Parallel Computing Center at the University of Liverpool.

I have been fortunate to have an opportunity to participate in ARTMIP project. I

want to express my gratitude to the lead members of ARTMIP for giving me a chance to participate in such an important international project, and for the collaboration.

Living in the UK and the US has been an unforgettable experience. Many thanks to the Howard family, especially Laura and Rod for their friendship, hospitality, and kindness for all these years (GDTBA!).

Finally, I cannot miss thanking my brother, Michał, who has been not only my first mathematics, physics, and computer science teacher, but also my science hero. Words cannot express my gratitude to him. I am also grateful to my parents, Olga and Wojciech for their love, never-ending support, and generosity over all these years. The moral values and an ethical code that I have been taught help me become a better person every day.

Last but not least, I would like to thank Dr Andrzej Brzoza and Dr Jan Spaliński for their encouragement since the beginning of PhD studies.

Abstract

Weather phenomena have long been objects of studies in atmospheric and climate science research. Studies on weather phenomena incorporate meteorological data, climate model simulations, and knowledge of physical processes of the Earth's atmosphere. Understanding of the developing mechanisms, life cycles, and spatiotemporal dependencies of these phenomena requires accurately identifying them in space and time. Moreover, identifying weather phenomena in large-scale climate model simulations is critical for advancing our understanding of the Earth's climate and risks of future extreme weather events.

The main goal of this thesis is to design and develop pattern recognition methods that directly learn from examples of weather phenomena in climate data, rather than following heuristic algorithms containing threshold requirements on physical variables. In particular, we aim to classify and localise atmospheric river and blocking phenomena in global climate simulations and reanalysis data.

In this thesis, we propose a novel pattern recognition method for identifying atmospheric river phenomena in climate datasets. This method consists of topological data analysis and machine learning methods. We demonstrate that the proposed method is reliable, robust, and achieves high accuracy. Also, we test the method on a wide range of spatial and temporal resolutions of global climate model outputs. We find that the method achieves the highest classification accuracy for low-resolution climate model outputs.

Moreover, we propose a hierarchical pattern recognition method for identifying of atmospheric blocking phenomena in climate reanalysis products. This pattern recognition method is based on deep convolutional neural networks. We demonstrate that the proposed method accurately detects and localises atmospheric blocks in climate reanalysis data. We also find that the method achieves higher accuracy for classification and lower estimation error for localisation of blocking phenomena in regions of the Northern Hemisphere than in regions of the Southern Hemisphere.

Research outcomes presented in this thesis show that the proposed pattern recognition methods can be complementary tools to the existing identification methods of atmospheric rivers and blocks in climate data. In addition to that, the proposed methods offer automatic post-processing, quantitative assessment of climate datasets, and can facilitate analysis of the local impacts of weather phenomena on specific geographical areas.

Contents

List of Figures

List of Tables

Nomenclature

AB	Atmospheric Blocking
AR	Atmospheric River
CCC	Concordance Correlation Coefficient
CNN	Convolutional Neural Network
DL	Deep Learning
F1	F1 score
IWV	Integrated Water Vapour
ML	Machine Learning
MPE	Mean Percentage Error
NH	Northern Hemisphere
SH	Southern Hemisphere
SVM	Support Vector Machine
TDA	Topological Data Analysis
TMQ	Total Mass Quantity of water vapour content
VGG	Visual Geometry Group deep learning architecture

Chapter 1

Introduction

In this chapter, we provide an overview of the Earth's changing climate and weather phenomena (i.e., atmospheric rivers and atmospheric blocks) and the associated extreme weather events. We also describe the identification of weather phenomena and the most commonly used climate data representations. Next, we explain our motivation to study weather phenomena (here, atmospheric rivers and blocks) that often lead to extreme weather events. Furthermore, we formulate research questions on the identification of weather phenomena in climate data, and we describe the contributions of this thesis. Finally, we indicate a part of the published research work that was adapted to this thesis, and we outline the thesis structure.

1.1 Impact of Changing Climate on Weather Phenomena

The Earth's climate changes because of its natural variability and external forcings (e.g., solar radiation). Since the start of the 20th century, changes in the climatological normal have been linked to anthropogenic-induced increases in greenhouse gases (i.a., carbon dioxide, methane, and nitrous oxide) [5, 6, 7]. However, studies on rising concentrations of carbon dioxide in the atmosphere related to surface air temperature increases on the Earth have been dated back to as early as the late 19th century. A well-known example is Arrhenius's studies [8]. In his work, he noted that the surface air temperature variability could be significantly larger at the high latitudes of the globe because of an elevated level of concentration of carbon dioxide in the atmosphere.

Moreover, the aforementioned studies take account of the fact that we distinguish five

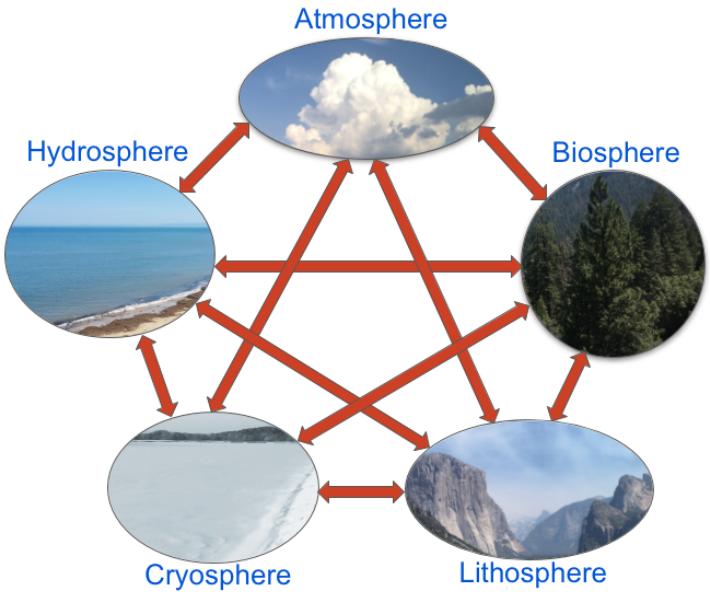
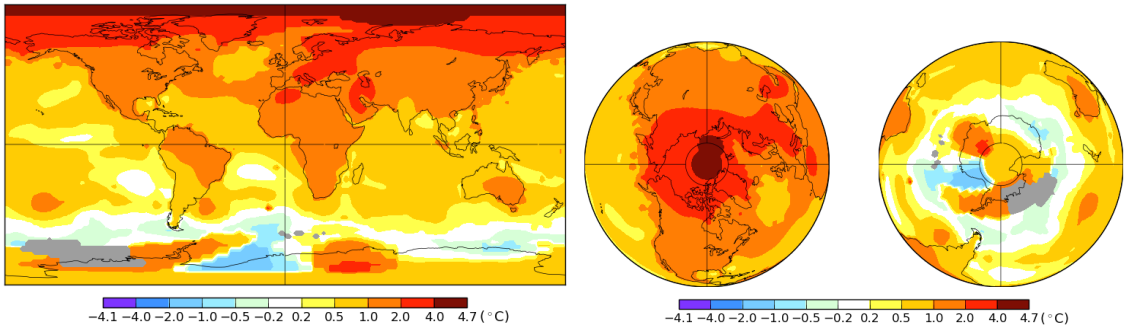


Figure 1.1: Diagram depicting the interactions of five major components of an idealised climate system on the Earth. Red arrows indicate two-way interactions between each component of the climate system.

major components of the Earth's climate system (i.e., the atmosphere, the hydrosphere, the lithosphere, the cryosphere, and the biosphere) that interact with one another (see Figure 1.1). By way of example, variations in the atmospheric temperature influence ice coverage, which in turn impact both the atmosphere and oceans. Another example is changes in the sea-surface temperature that have substantial implications for the atmosphere and weather phenomena in the middle latitudes of the globe [9].

The aforesaid process identified by Arrhenius [8] appears to be the first formal recognition of what has come to be a warming climate in the Earth's polar regions. The continuous increase of greenhouse gases emissions has accelerated warming in the Arctic more than in the globe as a whole in the past few decades. Climate model simulations have been projecting widespread Arctic warming as the climate system response to increasing atmospheric greenhouse gases concentrations for over thirty years [10, 11, 12]. A warming phenomenon in the north polar region is called Arctic amplification that is an active research area [13]. The rapid heating of the Arctic may be causally linked to a change in the frequency and intensity of atmospheric phenomena, for example, extreme weather events in mid-latitude regions [14, 15].



(a) (b)

Figure 1.2: A global image of the Earth showing the change of annual surface air temperatures (°C) over the 50 years, 1969-2018: (a) the equirectangular map projection of the globe. (b) Left side: the polar orthographic map projection of the Northern Hemisphere. Right side: the polar orthographic map projection of the Southern Hemisphere. Note: grey areas signify missing data. Source: The National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies, `http://data.giss.nasa.gov/gistemp/maps/`.

Furthermore, studies of observational records and global climate models indicate that the Arctic has approximately warmed twice as much as the rest of the Northern Hemisphere over the past fifty years [16] (see Figure 1.2 (a); the left side of Figure 1.2 (b)). Based on coupled climate model projections, the Fifth Assessment Report of the Intergovernmental Panel on Climate Change estimates that large temperature variability in the Arctic continues to increase as more greenhouse gases accumulate in the atmosphere [17]. The Antarctic is the exception where warming over the Southern Hemisphere high latitudes has been modest in comparison to the Arctic [18] (see the right side of Figure 1.2 (b)).

Sea ice loss accompanies rapid and widespread warming in the Arctic region [17]. A greater area of the region becomes seasonally ice-free (i.e., low albedo) [19] and extra solar radiation is absorbed during summers in vast extents of open-water of the Arctic Ocean. The absorbed solar energy is transferred to the overlying atmosphere as heat [20]. Hence, additional latent energy released to the atmosphere can impact the evolution of weather phenomena in the Arctic and beyond.

Long-term climate simulation experiments suggests that warming in the Arctic causes an increase in atmospheric water vapour content. Taking into account a well established physical law (i.e., the Clausius-Clapeyron relation) that says for each one degree of Celsius of warming, saturated air contains $6 - 7\%$ more water content [21]. Thus, warmer air can

hold more moisture, which in turn can fuel more frequent and stronger weather phenomena, for example, heavy precipitation in the lower latitudes of the globe [22].

Decades of studies of the global warming linked to the Arctic amplification indicate variations in atmospheric features, including the altering of large scale atmospheric circulation patterns [23], shifting the winds of the jet stream [24], slower progression of planetary-scale waves (i.e., Rossby waves) [25], and strengthening of some weather phenomena (e.g., atmospheric blocks [26]). In particular, the variations in weather phenomena can affect the resulting weather extremes (e.g., droughts, heatwaves, floods) in the middle latitudes of the globe [27].

Although changes in the Arctic influence the weather in the mid-latitudes, the mid-latitude circulation can also impact the Arctic itself. For instance, it has been observed that changes in the Arctic, such as the loss of sea ice cover, are linked to the poleward transport of moisture from outside the high latitudes [28]. These intrusions of moist air and heat fluxes are connected to weather phenomena, for instance, atmospheric rivers [29] and atmospheric blocks that are connected with Rossby wave breaking events [30].

Through observed changes in the key atmospheric features, there is a need for extensive studies on the bi-directional nature of Arctic amplification and the mid-latitude weather. That also suggests that the development of the ability for studying the spatiotemporal characteristics of weather phenomena under global warming is critical for advancing our knowledge of the Earth's changing climate, risk management, and informing governmental policy decisions.

1.2 Weather Phenomena and Extreme Events

Weather phenomena have long been an object of studies because they can bring both a severe risk and benefit to human beings [31] and the natural infrastructure [32]. The first discussions explaining the natural causes of specific types of weather phenomena have been started in ancient times [33]. Since that time, many scientists have made contributions to a variety of studies from data collection of meteorological observations to theories explaining the formation of weather phenomena. To give just one of many examples, contemporary studies incorporate in situ data collected over many years and from short-term numerical weather predictions to long-term climate modelling [34, 35, 36]. Also, the understanding of the main components of the climate system plays an important role in characterising dynamical processes in the atmosphere that can affect weather phenomena [37].

Formally, weather phenomena are also known as weather systems. They are mainly the result of complex dynamical processes happening in the lower atmosphere (i.e., the troposphere), variations in the large-scale atmospheric circulation patterns, and forcing from the sea-surface temperature. All weather phenomena vary on a range of timescales, from minutes or hours to a week or months, and a spatial extent (e.g., from a few kilometres to hundreds of kilometres). That is why the number of phenomena is large, and it is difficult to cover all of them.

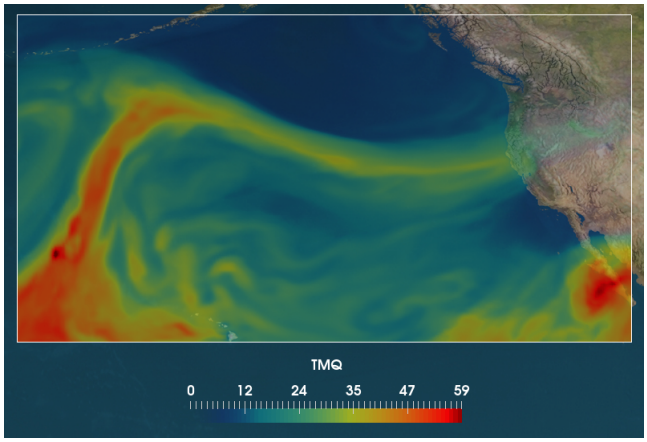
In this thesis, we follow one of the definitions of weather phenomena proposed by the American Meteorological Society (AMS) [38]: *"A typically transitory mesoscale to synoptic-scale region of atmospheric conditions representing a disruption or disturbance of the mean flow, most often used to describe areas of significant (but not necessarily severe or unusual) weather"*. However, the AMS definition is general and qualitative, it can be extended and translated into a quantitative one.

From here on out, we focus on two specific weather phenomena, atmospheric rivers and blocks, and we below provide their AMS definitions. We also describe extreme weather events that are often associated with both phenomena.

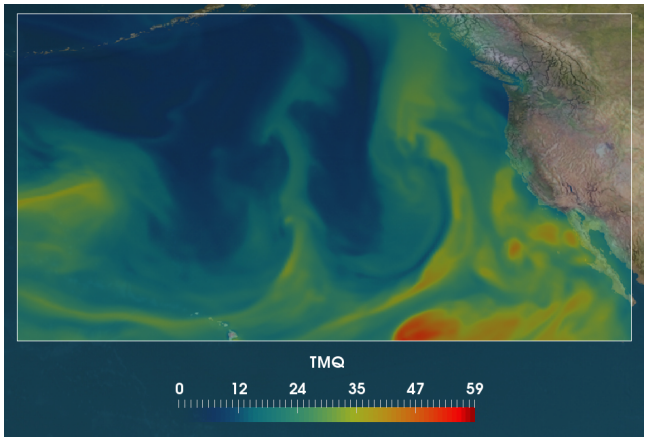
1.2.1 Atmospheric River Phenomena

Atmospheric Rivers (ARs) are dynamic filamentary structures of highly concentrated water vapour content in the atmosphere [39]. Figure 1.3 gives examples of an AR phenomenon and a non-AR situation. It displays the total column integrated water vapour variable (here, TMQ – the Total Mass Quantity of atmosphere water vapor content) in $kg\,m^{-2}$. AR phenomena can account for approximately 90% of poleward moisture transport and latent heat outside of the tropics from low to middle latitudes to landmasses [40]. These narrow synoptic-scale weather phenomena occupy approximately 10% of the available geographical longitude [41, 42]. ARs are usually formed over oceans and in the warm sectors of extratropical cyclones [43].

The American Meteorological Society glossary proposes the following definition of ARs [44]: *"A long, narrow, and transient corridor of strong horizontal water vapor transport that is typically associated with a low-level jet stream ahead of the cold front of an extratropical cyclone. The water vapor in atmospheric rivers is supplied by tropical and/or extratropical moisture sources. Atmospheric rivers frequently lead to heavy precipitation where they are forced upward - for example, by mountains or by ascent in the warm conveyor belt. Hor-*



(a)



(b)

Figure 1.3: (a) An example of an atmospheric river (i.e., a continuous narrow structure stretched from the Hawaiian Islands to the Pacific coast of North America); (b) An example of a non atmospheric river phenomenon on the Pacific coast of North America. Shown is the total column integrated water vapour variable (here, TMQ - the Total Mass Quantity of atmosphere water vapor content) in $kg\,m^{-2}$. The TMQ variable is from a simulation of the Community Atmosphere Model v5.1 at 100 km horizontal resolution [1].

izontal water vapor transport in the midlatitudes occurs primarily in atmospheric rivers and is focused in the lower troposphere". This AMS definition is the result of a consensus on a qualitative definition of ARs by the climate science community [45].

In spite of the fact that ARs can lead to extreme events (e.g., heavy precipitation and flooding or landslides) [46, 47], they also provide benefits to regions along mid-latitude

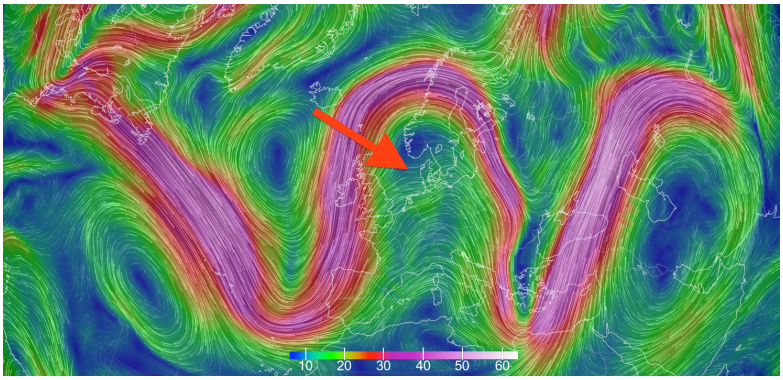
western continental boundaries. They contribute substantially to mountain snowpacks and provide $30 - 50\%$ of annual water supply to the western continental regions [43, 46]. That is why ARs are so crucial in the global hydrological cycle. ARs presence mostly supports hydrological balance in areas, such as North America (i.e. the western US) [48], Europe (the Atlantic European coast) [49], the Andes in South America [50], and South Africa [51]. Furthermore, the understanding of how ARs may change in a warmer climate is essential for prediction of frequency and intensity of precipitation under different global warming scenarios [52, 53].

1.2.2 Atmospheric Blocking Phenomena

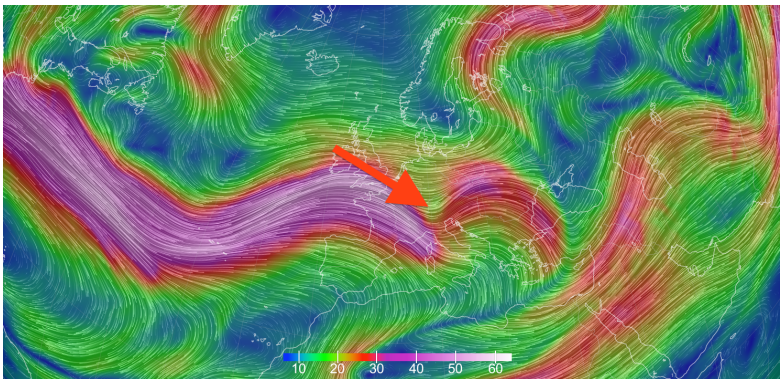
Atmospheric blocks (ABs) are persistent and quasistationary phenomena that reverse westerly flow at middle latitudes for an extended time [54, 55]. It means that these synoptic-scale weather phenomena can persist from days to several weeks [56]. Figure 1.4 shows examples of an AB phenomenon and a non-AB situation. It also displays the wind speed field in $m\ s^{-1}$ at 250 hPa pressure level. ABs can lead to many kinds of extreme weather events, such as heat waves [57], cold spells [58], droughts [59], and floods [60]. All these extreme events frequently occur in the middle latitudes, and they can have disastrous impacts on society [57], the natural infrastructure and the economy [61].

In the second half of the 20th century, many theoretical studies started looking for a physics-based model of AB dynamic development that could account for its spatiotemporal characteristics. Nevertheless, there is neither a theoretical model of ABs that can constitute for all of its observational characteristics nor there is a commonly accepted structural definition of these phenomena by the climate science community [62].

The American Meteorological Society glossary proposes the following definition of AB phenomena [63]: *"A blocking situation is attended by pronounced meridional flow in the upper levels, often comprising one or more closed anticyclonic circulations at high latitudes and cyclonic circulations at low latitudes (cut-off highs and cut-off lows). This anomalous circulation pattern (the "block") typically remains nearly stationary or moves slowly westward, and persists for a week or more. Prolonged blocking in the Northern Hemisphere occurs most frequently in the spring over the eastern North Atlantic and eastern North Pacific regions"*. The AMS definition is qualitative and provides general criteria for classifying a flow pattern as a blocked one. But it is still common to use regional criteria based on decades of observations to define AB phenomena.



(a)



(b)

Figure 1.4: (a) An example of atmospheric blocking (i.e., a high pressure pattern that blocks and diverts the jet stream). Here, it is shown a particular example of the Omega block that forms a shape resembling the Greek letter omega; (b) An example non atmospheric blocking phenomenon that does not cause unusual changes in the normal flow of the westerly winds; Shown is the wind speed field ($m\,s^{-1}$) at $250\,hPa$ pressure level (i.e., the elevation) from a simulation of the Global Forecast System produced by the National Centers for Environmental Prediction. Source: The US National Weather Service data visualised by `http://earth.nullschool.net`.

Since AB phenomena affect the weather of large parts of the mid-latitudes, they have progressively emerged as a key weather system to measure the ability of atmospheric models to forecast the weather. At many weather and climate forecast centres, ABs constitute an ideal benchmark to evaluate the ability of numerical models to forecast weather from short-term to long-term ranges and to represent the current climate [62].

1.2.3 Extreme Weather Events

Extreme weather events are usually caused by abnormal atmospheric phenomena and are also a natural part of the climate system. They often lead to high-impact weather events, such as heat waves [64], droughts [59], heavy precipitation [65], floods [60], cold spells [66], and low-quality air conditions [67]. As there is a long list of extreme weather events, the number of their definitions is vast. Furthermore, it is known that what is considered as an extreme event is always based on previous experience. The extensive review of all extreme events is out of the scope of this thesis.

In this thesis, we consider the most general definition of extreme weather events proposed by glossary of the Fifth assessment report of the Intergovernmental Panel on Climate Change (IPCC) [68]: *"An extreme weather event is an event that is rare at a particular place and time of year. Definitions of rare vary, but an extreme weather event would normally be as rare as or rarer than the 10th or 90th percentile of a probability density function estimated from observations. By definition, the characteristics of what is called extreme weather may vary from place to place in an absolute sense"*. Because of the nature of these events, some researchers often relax the definition of weather extremes. The relaxation usually allows climate scientists to add the level of intensity and temporal information to the definition. That can be used to define extreme climate events. The IPCC glossary proposes the following definition of extreme climate events [68]: *"When a pattern of extreme weather persists for some time, such as a season, it may be classed as an extreme climate event, especially if it yields an average or total that is itself extreme (e.g., drought or heavy rainfall over a season)"*.

1.3 Identification of Weather Phenomena in Climate Data

The identification of weather phenomena is fundamental to studies of atmospheric conditions in meteorology, atmospheric science, and climate science [69, 70]. First attempts at finding weather phenomena investigated observational data (often heterogeneous data [71, 72]) and ad hoc visual inspections by individuals, thus were limited in scope. However, the situation was transformed by the revolution of numerical weather predictions [34] and climate modelling [35, 36]. Furthermore, the process of climate data collection has been improved through the established global observational systems, remote sensing, and increased computational power [73, 74]. Nowadays, computational advancements allow

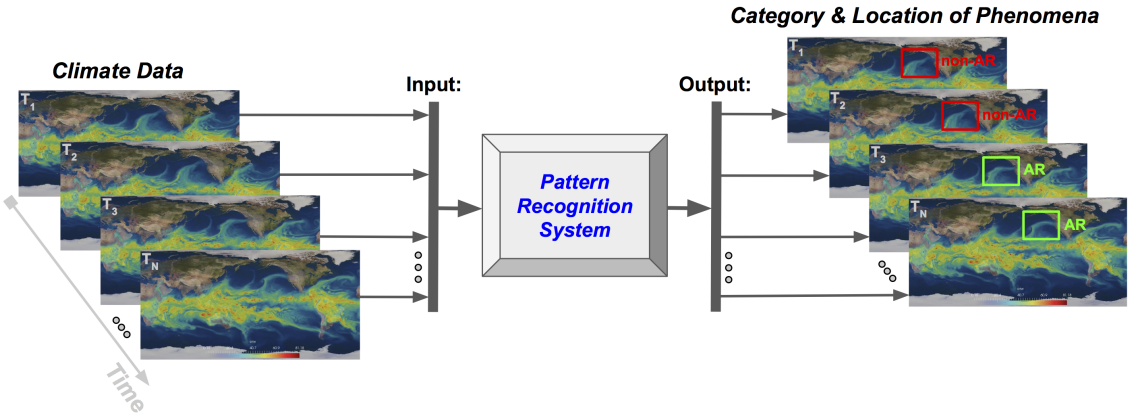


Figure 1.5: The scheme illustrates an example of a pattern recognition system for identifying weather phenomena (e.g., atmospheric rivers). The system takes input climate data or fields (i.e., physical variables laid out on the regular grid) at particular time steps $T_1 \dots T_N$, where N is the number of time steps (e.g., three hourly). Then, the pattern recognition system performs a recognition task on the input data. Next, the system outputs a category or a class of particular phenomena (e.g., atmospheric rivers (ARs) and non-atmospheric rivers (non-ARs)) and a location of the phenomena on the globe. The location can be segmentation masks (i.e., binary masks) or bounding boxes containing instances of specific phenomena.

researchers to process a deluge of data that can be assimilated into numerical weather predictions, and complex modelling of the Earth's climate system [75, 76].

Over the years, there has been a growth in the number of methods used to identify weather phenomena, such as tropical cyclones [77], hurricanes [78], atmospheric fronts [79], atmospheric blocks [80], and atmospheric rivers [2]. By far the best illustration of this is a typical pattern recognition system that is built from weather identification methods, as shown in Figure 1.5.

In the last few years, there has been growing attention to uncertainty in weather phenomenon detection methods. That may impact our understanding of future climate changes. For this reason, a few inter-comparison projects have been created to assess the uncertainty through collaborative experiments, for example, the Intercomparison of Mid-Latitude Storm Diagnostics (IMILAST) [77] and the Atmospheric River Tracking Method Intercomparison Project (ARTMIP) [81]. That uncertainty arises because there do not exist complete theoretical and quantitative definitions of weather phenomena, for instance,

atmospheric blocks [62]. In order to do quantitative research related to these weather phenomena, the climate science community has had to form detection methods to recognise specific phenomena independently. That is why most of the existing detection methods in the literature are predominantly heuristic [82]. In other words, these methods consist of a set of rules used to isolate specific weather phenomena in climate fields.

In atmospheric science, meteorology, and climate science research, one can distinguish three main categories of climate data representation: model simulation outputs, observational datasets (e.g., satellite images and weather network stations), and reanalysis products (i.e., generated by combining climate models with observations). Climate models are based on the fundamental laws of physics governing the climate system. They can produce dynamic, chaotic behaviour that is characteristic of the climate system. Moreover, climate models produce long-term climate studies of the Earth's atmosphere [35, 83]. Climate reanalysis products are hybrid model-observation based datasets. They are produced by advanced data assimilation methods applied to observational data [75, 76]. The observations are assimilated into highly constrained climate model simulations. The reanalysis is a comprehensive historical record of the Earth's climate and provides a reliable way to monitor how fast it changes.

Both climate models and reanalysis products often come in the form of fields that are laid out on the regular grid. They are are sampled at daily or hourly time steps (e.g., 3-hourly and 6-hourly resolution). The typical horizontal resolutions are approximately at $25\,km$, $50\,km$, and $100\,km$ (i.e., thousands of grid points), and the data are sampled at many pressure levels (e.g., 60 elevations).

1.4 Motivation

This thesis is at the intersection of computer science and climate science research. Our motivation for studying weather phenomena is two-fold.

Firstly, we consider this study from the point of view of research on climate science. An investigation of the climatological statistics of atmospheric rivers (ARs) and atmospheric blocks (ABs) can provide an understanding of how these phenomena may vary under various climate change scenarios. Because both ARs and ABs play an important role in the middle latitude climate and weather, it is critical to produce and analyse the statistics, such as the frequency, location, and intensity of these phenomena. Furthermore, such valuable statistical information may be useful for the general public, policymakers,

and stakeholders for risk assessment of extreme events.

Secondly, we motivate our study from the computer science perspective and the need for new data analysis methods for climate science applications. Climate model simulations are one of the primary means for understanding the past climate and making climate predictions. Over the decades of climate modelling research, the models have become more complex and produce an ever-increasing amount of data from simulations or reanalysis products. The produced large datasets demand post-processing and quantitative analyses. A manual identification of weather phenomena (e.g., ABs and ARs) is a labour-intensive process requiring analysis of thousands of individual snapshots of climate model simulations. That provides a strong motivation for developing accurate, fully automated methods when considering the analysis of large climate datasets.

1.5 Research Questions and Contributions

This section considers the research questions that this thesis seeks to answer and itemises the main contributions of the work.

1.5.1 Research Questions

In this thesis, we present three research questions, each question consists of three subquestions. In particular, each research question is related to the identification task of weather phenomena in global climate datasets and the effectiveness of proposed weather pattern recognition methods. We focus on identifying Atmospheric River (AR) and Atmospheric Block (AB) phenomena, and we formulate the following research questions:

1. Can pattern recognition methods automatically discover regularities in raw climate data to identify AR and AB phenomena? (RQ1)

    a) Do pattern recognition methods based on machine learning models facilitate the characterisation of weather phenomena? (RQ1.a)

    b) To what extent can pattern recognition methods replace weather phenomenon identification methods? (RQ1.b)

    c) Can machine learning models provide the climatological statistics of weather phenomena from global climate model outputs? (RQ1.c)

2. Is it possible to detect ARs by extracting high-level representations of AR phenomena from climate data? (RQ2)

    a) How topological features of ARs can contribute to the characterisation of the structure of ARs? (RQ2.a)

    b) Do machine learning models benefit from topological features to learn high-level representations of ARs from climate model outputs? (RQ2.b)

    c) Are AR pattern recognition methods sensitive to changes in spatiotemporal resolution of climate model simulation outputs? (RQ2.c)

3. Is it possible to detect ABs by learning high-level representations of AB phenomena from multivariate and high-dimensional climate reanalysis data? (RQ3)

    a) Can deep learning models automatically extract high-level representations of ABs to detect these phenomena in climate reanalysis products? (RQ3.a)

    b) Is it possible to predict the localisation and geometric properties of ABs in climate reanalysis data? (RQ3.b)

    c) To what extent can the depth of deep learning architectures contribute to the AB detection performance in various geographical regions over the globe? (RQ3.c)

1.5.2 Contributions

The contributions of this thesis are as follows:

- We show that the proposed weather pattern recognition methods can be complementary to heuristic identification methods of weather phenomena when a vast amount of labelled climate data is available (RQ1.a).

- We demonstrate that machine learning models can learn the informative characterisation of weather phenomena that is reflected in high recognition performance (RQ1.b).

- We show that machine learning models can accurately compute the climatological statistics of weather phenomena from global climate model outputs and reanalysis products (RQ1.c).

- We are first to quantitatively demonstrate that topological features of ARs can contribute to the characterisation of the structure of ARs. We also show that machine learning models benefit from topological features in learning a high-level representation of ARs from climate simulation outputs (RQ2.a, RQ2.b).

- We carry out computational experiments that show that AR recognition methods' performance is sensitive to changes in spatiotemporal resolution of climate model simulation outputs (RQ2.c).

- We are the first to find that deep learning models can automatically extract high-level representations of ABs from data and then detect ABs in climate reanalysis products. Moreover, we show that it is possible to accurately predict the localisation and geometric properties of ABs (RQ3.a, RQ3.b).

- We carry out analyses that demonstrate that the detection performance of ABs significantly increases when using deep CNN architectures, while the estimation error of the phenomena localisation significantly decreases for the shallow ones. Also, the proposed CNN architectures tend to achieve higher performances for detecting and localising of ABs in regions of the Northern Hemisphere than in regions of the Southern Hemisphere (RQ3.c).

These contributions were published in peer-reviewed scientific journals and peer-reviewed conference proceedings (see Subsection 1.6), and involved collaboration, including joint work with: Dr Prabhat Ram (NERSC, Lawrence Berkeley National Laboratory, USA), Dr Karthik Kashinath (NERSC, Lawrence Berkeley National Laboratory, USA), Dr Dmitriy Morozov (CRD, Lawrence Berkeley National Laboratory, USA), Dr Jan Balewski (NERSC, Lawrence Berkeley National Laboratory, USA), Dr Christine Shields (Climate and Global Dynamics Division, NCAR, USA), Dr Jonathan Rutz (National Weather Service Western Region Headquarters, NOAA, USA).

1.6 Co-Authored Papers

This thesis is written based on the following co-authored publications in peer-reviewed scientific journals and conference proceedings. The content of Chapter 3 has been published in the following journals and conference proceedings:

- **Muszynski, G.**, Kashinath, K., Kurlin, V., Wehner, M., Prabhat. *"Topological data analysis and machine learning for recognizing atmospheric river patterns in large climate datasets."*. Geoscientific Model Development, 2019 [1].

- **Muszynski, G.**, Kurlin, V., Morozov, D., Wehner, M., Kashinath, K., Prabhat. *"Topological Methods for Pattern Detection in Climate Data"*, Big Earth Data Analytics in Earth, Atmospheric and Ocean Sciences, AGU Books Wiley, 2020 (in press) [84].

- Shields, C. A., Rutz, J. J., Leung, L.-Y., Ralph, F. M., Wehner, M., Kawzenuk, B., Lora, J. M., McClenny, E., Osborne, T., Payne, A. E., Ullrich, P., Gershunov, A., Goldenson, N., Guan, B., Qian, Y., Ramos, A. M., Sarangi, C., Sellars, S., Gorodetskaya, I., Kashinath, K., Kurlin, V., Mahoney, K., **Muszynski, G.**, Pierce, R., Subramanian, A. C., Tome, R., Waliser, D., Walton, D., Wick, G., Wilson, A., Lavers, D., Prabhat, Collow, A., Krishnan, H., Magnusdottir, G., Nguyen, P. *"Atmospheric River Tracking Method Intercomparison Project (ARTMIP): project goals and experimental design."*, Geoscientific Model Development, 2018 [85].

- Rutz, J. J., Shields, C. A., Lora, J. M., Payne, A. E., Guan, B., Ullrich, P., O'Brien, T., Leung, L. R., Ralph, F. M., Wehner, M., Brands, S., Collow, A., Goldenson, N., Gorodetskaya, I., Griffith, H., Kashinath, K., Kawzenuk, B., Krishnan, H., Kurlin, V., Lavers, D., Magnusdottir, G., Mahoney, K., McClenny, K., **Muszynski, G.**, Nguyen, P. D., Prabhat, Qian, Y., Ramos, A. M., Sarangi, Ch., Sellars, S., Shulgina, T., Tome, R., Waliser, D., Walton, D., Wick, G., Wilson, A. M., Viale, M. *"The atmospheric river tracking method intercomparison project (ARTMIP): quantifying uncertainties in atmospheric river climatology."*, Journal of Geophysical Research: Atmospheres, 2020 [3].

The content of Chapter 4 has been published in conference proceedings:

- **Muszynski, G.**, Prabhat, Balewski, J., Kashinath, K., Wehner, M. and Kurlin, V. *"Atmospheric Blocking Pattern Recognition in Global Climate Model Simulation Data."*, Proceedings of the 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, IEEE, 2021 [4].

1.7 Thesis Structure

The outline of the rest of this thesis is, as follows:

- Chapter 2 presents a literature review and related work.
- Chapter 3 describes a project on AR recognition problem, results, and discussions.
- Chapter 4 describes a project on AB recognition problem, results, and discussions.
- Chapter 5 concludes the thesis and presents lessons learnt, response to the research questions, and future work.

Chapter 2

State of the Art in Weather Pattern Recognition

In this chapter, we review related work on pattern recognition for weather phenomena in climate data and describe the existing methods for identifying atmospheric rivers and atmospheric blocks. We also outline differences and similarities between the existing methods that have been developed by the climate science community. That is followed by examples of newly applied machine learning and deep learning detection methods to weather pattern recognition problems. Finally, we discuss the limitations of the existing research on pattern recognition for weather phenomena in climate data.

2.1 Weather Pattern Recognition in Climate Science

Recognizing patterns of weather phenomena in climate data has progressively emerged as one of the most active research areas in climate science over the last decade. Pattern recognition of weather phenomena aims at automatically identifying a wide range of atmospheric conditions (i.e., objects or contours) in climate data (i.e., images and fields from climate model outputs). With the advent of computer technology and the development of new image processing techniques, some of the first fully automated identification methods have been developed to perform detection of weather phenomena, such as tropical cyclones [86, 87], extratropical cyclones [88, 89], weather fronts [90, 91], and the jet stream [92, 93]. These pioneering studies showed the great potential of pattern recognition methods in identifying weather phenomena for the following reasons [69]:

- The development of automated identification methods eliminates human subjectivity from the computations and detection procedures;
- Automated identification methods can be applied to global climate datasets and to determine the number of occurrences of specific weather phenomena (i.e., their spatial distributions);
- These methods can be used in validating specific aspects of climate simulations, for instance, the ability of climate models to characterise features or physical structure of weather phenomena;
- Pattern recognition methods can provide new insights into the understanding of weather phenomena under consideration by the automatic discovery of regularities in data.

For example, it has been observed that advances in the development of pattern recognition methods helped substantially discover the linkage between cyclones and fronts and extreme precipitation [64, 94, 95].

In general, weather phenomenon identification methods are based on the Eulerian feature diagnostics (i.e., condition-based algorithms) and the Lagrangian flow feature trajectory techniques (i.e., tracking-based algorithms) [82, 96]. The Eulerian diagnostics use typically two-dimensional input fields from climate model outputs. The diagnostics identify specific weather phenomena through the application of one, or, a couple of thresholds, and geometric criteria to the climate fields. The Lagrangian techniques often use three-dimensional fields from climate model outputs (i.e., the two-dimensional climate fields, plus the time axis) as an input. The techniques calculate flow feature trajectories, and then they apply selection criteria to identify the desired weather phenomena. The Eulerian and Lagrangian-based methods yield a two or three-dimensional binary mask with the value one at grid points that meet the condition or the criteria, and the value zero at grid points that do not.

Most existing weather phenomenon identification methods are engineered heuristics based on human expertise and domain knowledge. That means that specific threshold conditions, or criteria, used to define certain atmospheric phenomena are imposed on the Eulerian and Lagrangian algorithms, which are the core of the methods. The threshold conditions are often observationally derived values from satellite data or aircraft reconnaissance. Different identification methods are built upon various hard constraints on

thresholds of relevant physical variables or fields, such as surface pressure, temperature, and wind speed. For all these reasons, designing an objective identification method is a difficult task. It requires setting up various sets of thresholds or criteria to characterise weather phenomena that are hard to be justified. That is why there have been a few climate science community efforts to intercompare different identification methods of weather phenomena, for example, a project on extratropical cyclones diagnostics—the Intercomparison of Mid-Latitude Storm Diagnostics (IMILAST) [77]. Moreover, it often happens that weather phenomena neither have a unique theoretical model nor have a clear empirical definition that is universally accepted by the climate science community. For instance, weather phenomena can have various spatial structures or shapes that depend on a specific geographic location on the globe and the seasonality.

There are many ongoing efforts to search for alternative pattern recognition methods for weather phenomena other than conventional identification methods based on threshold conditions. These methods include the discovery of coherent structures in spatiotemporal data [97], a supervised and semi-supervised machine learning methods along with deep learning techniques [98, 99, 100]. The alternative methods by their inherent design circumvent a critical selection of suitable thresholds of different physical variables to characterise weather phenomena. For example, deep learning-based methods can learn robust representations of images from raw pixels, outperforming handcrafted features [101]. Despite deep learning success in the computer vision field, the adoption of deep neural networks in climate science research is still unexplored.

2.2 Atmospheric River Pattern Recognition Methods

Atmospheric River (AR) phenomena have been first introduced in the scientific literature by Newell et al. [39]. In that research work, the phenomena were called “rivers”, because they are long and narrow streams of moisture moving at high speeds in the lower troposphere. Moreover, some ARs are capable of carrying as much water as the Amazon river. However, the term “Atmospheric River” has been first proposed by Zhu et al. [41]. Those pioneering studies of Newell and Zhu examined a daily global tropospheric water vapour flux values or Integrated Water Vapour Transport (IVT) to indicate the presence of these weather phenomena in global observational datasets.

In the decades since the seminal work of Newell and Zhu, many researchers were seeking to improve the understanding of ARs and their role in climate and weather. A first

identification method has been developed by the same authors of the earlier studies [42]. In order to detect ARs, the method uses the key characteristic of AR phenomena that is the magnitude of the moisture fluxes in the troposphere. Nevertheless, the authors' approach was revolutionary; the moisture flux was difficult to observe directly. Hence, many complementary methods have been developed to detect ARs in various meteorological and climate datasets over the past two decades.

Table 2.1 and Table 2.2 summarise selected previous studies on AR identification methods on climate datasets, such as climate model outputs, reanalysis products, and observational data. The listed studies are examples of most representative AR identification methods in use by the climate community. It is difficult to describe each approach in detail; that is why we group methods concerning a temporal criterion. Table 2.1 lists condition-based methods that are independent of a time dimension. They identify ARs in single time slices of spatial climate fields where specific conditions are met in the Eulerian sense. On the other hand, Table 2.2 shows tracking-based methods that are dependent on the time dimension. These methods detect ARs and follow the phenomena in time and space in the Lagrangian sense. Each study (i.e., *Authors* columns) is listed by row and refers to an individual or group who developed the method. Methods identifiers (i.e., *Method ID*) are used in Chapter 3 and Chapter 4. The *Threshold requirements* in the third columns refer to a choice of either the absolute threshold or the relative threshold that must be satisfied by objects to be identified as AR phenomena. The last column (i.e., *Region*) refers to whether or not a method detects ARs globally or locally in climate data.

The first major difference between methods is in how to choose a characteristic variable of ARs on which to perform thresholding. It is usually a choice between Integrated Water Vapour (IWV) and IVT. IWV is one of the most important variables in use because of the pioneering study of Ralph et al. [43]. This variable is readily available from satellite data and aircraft observations over the oceans. Nowadays, many methods have been shifting toward IVT that is a product of on zonal and meridional wind data, and specific humidity variables [109]. Furthermore, IVT allows to do AR analysis over the land and to predict high-impact hydrological events [123] where IWV may not be available.

Another major area where methods differ is whether the threshold can be absolute (i.e., held constant), or be relative (i.e., changing as a function of, for example, location). A wide range of AR identification methods follow studies of Ralph et al. [43] and Rutz et al. [109] that use an absolute threshold value (i.e., observationally derived value from satellite data and aircraft reconnaissance) to detect ARs [81, 103, 107, 110, 112, 114, 115, 116, 120].

Authors (Method ID)	Threshold requirements	Region
Brands et al. [102] (C1)	Relative IVT and Absolute IVT	Atlantic Coast of North Africa, Atlantic Coast of Europe, Pacific Coast of North America
Goldenson et al. [103] (C2)	Absolute IWV	Pacific Coast of North America
Guan and Waliser [104] (C3)	Relative IVT and Absolute IVT	Global
Lora et al. [105] (C4)	Relative IVT	Global and Pacific Coast of North America
Mundhenk et al. [106] (C5)	Relative IVT	Global
Hagos et al. [107] (C6)	Absolute IWV	Pacific Coast of North America
Lavers et al. [108] (C7)	Relative IWV	Atlantic Coast of Europe
Rutz et al. [109] (C8)	Absolute IVT	Global
Ralph et al. [43] (C9)	Absolute IWV	Pacific Coast of North America
Shields and Kiehl [81, 110] (C10)	Relative IWV	Pacific Coast of North America, Atlantic Coast of Europe
Prabhat et al. [111, 112] (C11)	Absolute IWV	Global
Liu et al. [113] (C12)	No thresholds	Global

Table 2.1: The table lists previous studies on condition-based identification methods of atmospheric rivers. IVT is integrated water vapour transport, and IWV stands for integrated water vapour column.

Authors (Method ID)	Threshold requirements	Region
Sellars et al. [114] (T1)	Absolute IVT	Global
Zhu et al. [42] (T2)	Relative IVT	Global
Gershunov et al. [115] (T3)	Absolute IWV	Pacific Coast of North America
Leung and Qian [116] (T4)	Absolute IVT	Pacific Coast of North America
Payne et al. [117, 118] (T5)	Relative IVT and Absolute IWV	Pacific Coast of North America
Ramos et al. [119] (T6)	Relative IVT	Atlantic Coast of Europe, Atlantic Coast of South Africa
Ullrich et al. [120] (T7)	Absolute IVT	Global
Huang et al. [121] (T8)	Relative IVT	Pacific Coast of North America
Racah et al. [122] (T9)	No thresholds	Global
Mudigonda et al. [100] (T10)	No thresholds	Global

Table 2.2: The table lists previous work on track-based identification methods of atmospheric rivers. IVT is integrated water vapour transport, and IWV stands for integrated water vapour column.

The threshold value observed in these studies is an IWV value $\geq 20\, kg\, m^{-2}$ or IVT value $\geq 250\, kg\, m^{-1}\, s^{-1}$. Other methods use a relative threshold value based on a statistical approach and anomaly analysis [81, 105, 106, 108, 110, 119, 121]. For example, the approach proposed by Lavers et al. [108] utilises an IVT value $\geq$ the 85th percentile of seasonal or climatological IVT. A few methods take further steps and employ both the absolute thresholds and relative thresholds [102, 104, 117, 118].

Another set of methods is based on modern pattern recognition approaches, such as machine learning methods, including deep learning neural networks [101]. Deep neural networks, especially deep convolutional neural networks, have achieved state-of-the-art results in image classification tasks on many benchmark image datasets (e.g., ImageNet [124]) in recent years. The advantage of this set of methods is that they can identify ARs as spatial objects without relying on threshold requirements of IWV and IVT variables to characterise ARs. Those methods use statistical learning algorithms that automatically extract

high-level representations of AR phenomena from climate data and assign example (e.g., a cropped patch of climate field) to AR category or the other one (i.e., a non AR).

Liu et al. [113] were first to demonstrate that deep convolutional neural networks (CNNs) architecture could be successfully applied to predict the class label for AR phenomena. In that work, the authors performed the binary classification task on centred, cropped patches from two-dimensional and single time slices of IWV fields. Moreover, Mudigonda et el. [100] employed CNNs to a supervised segmentation task of ARs in climate model outputs. The authors applied the state-of-the-art CNN architecture, called Tiramisu [125], to semantic segmentation of climate model fields to classify every grid point in a field as a member of one of three classes (i.e., ARs, tropical cyclones, and background). The Tiramisu architecture is based on Densely Connected Convolutional Neural Networks (DenseNets) used for image classification task [126]. In that study, cropped patches from two-dimensional and single time slices of IWV fields were used as the input for the CNN architecture. Furthermore, Racah et al. [99] proposed a novel multichannel spatiotemporal CNN architecture for semi-supervised bounding box prediction of AR phenomena in climate model output. The authors applied the architecture for a multi-class detection and localisation of four weather phenomena: ARs, tropical cyclones, extra-tropical cyclones, and tropical depressions. The CNN architecture is based on a deep auto-encoding CNN architecture for bounding-box regression task (i.e., a bounding box location and the size, class associated with the bounding box, and the confidence level of the bounding box). In this study, it was demonstrated that the use of three-dimensional CNN encoder-decoder architecture could perform better than their two-dimensional counterparts. The input for the CNN architecture was the global multichannel climate model output (i.e., height, width, time), including IWV fields. The fields were split into a grid of square boxes.

As we can see in Table 2.1 and Table 2.2, the methods detect AR phenomena over specified regions of interest and on the global scale. Those methods that identify ARs over certain regions mostly focus on the Pacific coast of North America region and the Atlantic coast of the European region, for instance, [43, 81, 102, 103, 105, 107, 108, 115, 116, 119, 121]. There are also a couple of methods that detect ARs on the whole globe [99, 100, 104, 106, 109, 112, 113, 114, 120, 127]. Up until this point, there are only two methods that uniquely study the Atlantic coast of North Africa and South Africa [102, 119].

2.3 Atmospheric Blocking Pattern Recognition Methods

Atmospheric Block (AB) phenomena were first introduced in the scientific literature by Garriott in 1904 [128]. Nevertheless, the interest around ABs has been aroused with the pioneering observational studies of Berggren et al. [129], Rossby [56], and Rex [54, 55] in the middle of the last century. In particular, the authors of these studies were the first to propose theories accounting for the dynamics of AB onset, maintenance, and decay. They also introduced the term "Atmospheric Blocking" or "Blocking Action" for the phenomena. Because ABs are a resilient obstruction of the normal west-to-east atmospheric circulation in the middle latitudes of both hemispheres. In other words, ABs are able to block and temporally redirect the jet stream that has a pronounced effect on middle latitude weather and climate [130].

The pioneering work of Berggren, Rossby, and Rex have been the inspiration for numerous observational and theoretical studies on ABs [131]. For example, some of the theoretical studies were considering the formation of ABs and attempting to provide a universally accepted structural definition of these phenomena [132, 133, 134]. Many of the observational studies have been devoted to developing approaches that can be used to investigate the geographical distribution of ABs observational statistics. However, the lack of a universal AB phenomenon definition is illustrated by the variety of identification methods developed over decades.

Table 2.3 and Table 2.4 summarise selected previous studies on visual inspection procedures of ABs and objective identification methods of these phenomena, respectively. The studies were conducted on climate datasets, such as observational data, climate model outputs, and reanalysis products. The listed studies are examples of the most representative approaches of identifying ABs, but these are not exhaustive lists of approaches in use by the climate science community. Each study (i.e., *Authors* column) is listed by row and refers to an individual or group who developed the approach or method. The *Variable requirements* in the second column refers to the choice of physical variable or data collection (e.g., geopotential height). The third column (i.e., *Criteria*) contains information on whether the authors followed specific criteria to characterise AB phenomena in developing their inspection procedure. The last column (i.e., *Region*) refers to whether or not the procedure was developed for regions in the Northern Hemisphere or the Southern Hemisphere. Furthermore, Table 2.4 lists objective identification methods that employ a combination of the Eulerian diagnostics and Lagrangian tracking techniques. In that table, the *Type of*

Authors	Variable requirements	Criteria	Region
Elliott and Smith [135]	Geopotential height	Customised criteria	Northern Hemisphere
White and Clark [136]	Geopotential height and Mean sea level pressure	Modified Rex's criteria	Northern Hemisphere
Austin [137]	Geopotential height and Mean sea level pressure	Rex's criteria	Northern Hemisphere
Treidl et al. [138]	Geopotential height and Mean sea level pressure	Modified Rex's criteria	Northern Hemisphere
Dole [139]	Geopotential height	Modified Rex's criteria	Northern Hemisphere
Charney et al. [140]	Geopotential height	Modified Rex's criteria	Northern Hemisphere
van Loon [141]	Geopotential height	Rex's criteria	Southern Hemisphere
Taljaard [142]	Geopotential height	Rex's criteria	Southern Hemisphere
Wright [143]	Geopotential height and Mean sea level pressure	Rex's criteria	Southern Hemisphere

Table 2.3: The summary of previous studies on visual identification procedures of atmospheric blocking phenomena.

method refers to whether the method was applied to a single time-varying central latitude (1D) or the full latitude-longitude (2D) range.

Since the seminal studies of Rossby [56] and Rex [54, 55], many researchers have sought to improve the scientific understanding of ABs role in climate and find a suitable approach for identification of those phenomena. The major difference between approaches is in whether to perform a visual inspection of synoptic weather maps in order to investigate examples of ABs or apply an objective method for identifying the phenomena. The first attempts of the climate science community have been concerned with visual inspection techniques of flow field patterns (i.e., synoptic weather maps) to indicate the presence of ABs [135, 136, 137, 138, 139, 141, 142, 143, 151]. The visual inspection technique requires

Authors	Variable requirements	Type of method	Region
Lejenäs and Økland [144]	Geopotential height	1D	Northern Hemisphere
Lejenäs [145]	Geopotential height	1D	Southern Hemisphere
Tibaldi and Molteni [146, 147]	Geopotential height	1D and 2D	Northern Hemisphere
Dole and Gordon [148]	Geopotential height	1D and 2D	Northern Hemisphere and Southern Hemisphere
Schwierz et al. [149]	Potential vorticity	2D	Northern Hemisphere and Southern Hemisphere
Pelly and Hoskins [150]	Potential vorticity and Potential temperature	2D	Northern Hemisphere and Southern Hemisphere
Pinheiro et al. [80]	Geopotential height and Potential vorticity	2D	Northern Hemisphere and Southern Hemisphere

Table 2.4: The summary of previous studies on objective identification methods of AB phenomena.

the weather maps that often consist of the plotted daily or monthly geopotential height and mean sea level pressure data.

Elliott and Smith [135] were the first to propose a way of visually inspecting historical weather maps in order to indicate the presence of ABs. In their work, they computed surface pressure anomalies from the climatological mean. Then, they used local geopotential height maxima (i.e., the anomalies exceeding a given threshold for a particular time) to indicate the presence of ABs. Moreover, the series of seminal works of Rex [54, 55] describes the well-established criteria to make the inspection procedure as objective as possible. In his works, Rex defined the following five criteria that synoptic situations with pressure patterns should have in order to be identified as ABs:

- The westerly stream splits into two separate branches and each branch transports significant atmospheric mass;
- This double-stream structure extends for at least 45 degrees of longitude;
- There is a sharp transition from zonal flow upstream to meridional flow downstream of the stream split;
- The above pattern persists for at least ten days.

Since Rex's publications, there has been an increasing interest in using his criteria in visual inspection procedures. White and Clark [136] used the altered version of Rex's criteria, which they applied to monthly mean synoptic weather maps. They used monthly mean geopotential height and mean sea level pressure to determine if AB phenomena were present for enough time to dominate the monthly mean pressure pattern. Treidl et al. [138] also modified Rex's criteria. They were able to indicate ABs if the surface pressure had northward spatial extent and the half of minimum duration established by Rex. However, Austin [137] used the same criteria as Rex proposed in his studies and investigated geopotential height and mean sea level pressure data to indicate AB situations. The main difference is that he focused on the role of planetary waves (i.e., wave numbers) in the presence of AB phenomena. Other research works, such as Dole [139] and Charney et al. [140], investigated whether particular statistical approaches could be useful in identifying ABs by inspecting the persistence of anomalies associated with blocking phenomena.

A common feature of all above studies is that they were carried out for sectors in the Northern Hemisphere. The first study of AB phenomena in the Southern Hemisphere was made by van Loon [141]. In the study, the author also used synoptic weather maps to determine the presence of ABs. Nevertheless, the blocking analysis was limited in scope because of data available for that region. Taljaard [142] and Wright [143] were the first to show notable frequency statistics of ABs in the Southern Hemisphere. In their research works, they used the same kind criteria to identify ABs as those criteria set down by Rex.

Visual identification procedures often lead different researchers to potentially inconclusive results [139, 140]. That is why much effort has been spent in the last 40 years to develop objective methods for identifying AB phenomena in global climate model simulations, reanalysis products, and observational measurements. To identify ABs, most of the methods translated Rex's criteria for AB phenomena into algorithms. These methods use an algorithmic approach that translates criteria of the kind proposed by Rex into the

automatic identification methods. Lejenäs and Økland [144] were the first to use Rex's criteria to determine where and when ABs occur by translating the criteria in the algorithmic procedure. In that study, they looked for a reversal in the geopotential height gradient in the middle latitude of the Northern Hemisphere. A similar approach was applied to the geopotential height fields in the Southern Hemisphere by Lejenäs [145]. Tibaldi and Molteni's method [146] is based on that proposed by Lejenäs and Økland. The method requires a minimum westerly flow north of the reversal in order to ensure that the main westerly stream was split around ABs. Next, the method was extended by adding a new condition to its previous version [147]. Dole and Gordon [148] proposed a method that searches for large anticyclonic anomalies that are mostly associated with AB phenomena. Pelly and Hoskins [150] developed a new method that searches for a reversal of the mean flow to determine the presence of AB phenomena. Following Pelly and Hoskins, Schwierz et al. [149] developed a novel method that searches for regions of persistent column-averaged negative potential vorticity anomalies. Recently, Pinheiro et al. [80] developed a new AB detection framework that combines three methods [146, 148, 149] in order to identify ABs. In that work, they employed each of the three methods and applied across the full latitude-longitude (2D) range of middle latitud regions in both the Northern Hemisphere and Southern Hemisphere.

2.4 Limitations of the Existing Research

Most AR and AB identification methods are engineered heuristics. The bottleneck of these methods is that important human expertise is required in design of methods for recognising specific atmospheric phenomena. That makes the identification methods strongly task-dependent. These identification methods also require setting up various sets of thresholds conditions and geometric criteria to characterise weather phenomena. That is not easy to uncover and justify. In addition, the non-uniformity in research on AR and AB phenomena usually results in discrepancies between outputs of different identification methods.

In contrast to engineered heuristic methods, there are machine learning and deep learning methods that can circumvent the task-dependency issue. These methods rely on the assumption that they can be trained on raw data. However, the disadvantage of deep learning methods is that a large number of training instances are required for training in comparison with heuristic approaches. Besides, there is a small amount of available labelled datasets. That significantly reduces the range of deep learning methods that can be

applied. Another issue is that many existing machine learning and deep learning methods do not take into account the fact that weather phenomena change over time.

This thesis attempts to overcome most of the aforementioned limitations. AR and AB identification methods will be designed to include invariant topological information and learn spatial dependencies by applying machine learning and deep learning methods. Research on AR and AB phenomena will be carried out on massive datasets of climate model outputs and reanalysis products. The labelled data for machine learning and deep learning models will be provided by the-state-of-the-art identification methods of AR and AB phenomena [112, 149].

Chapter 3

Atmospheric River Phenomenon Identification in Climate Model Outputs and Reanalysis Products

In this chapter, we mainly focus on data-driven approaches to characterise and to identify Atmospheric River (AR) phenomena [1]. We are interested in exploring alternative methods for identifying ARs, such as Topological Data Analysis (TDA) and Machine Learning (ML) methods for pattern recognition. These methods automatically extract underlying features of ARs from climate data and recognise AR phenomena. That corresponds to overcoming a critical challenge of selecting suitable threshold requirements (i.e., absolute or relative thresholds) of AR identification methods (see Chapter 2.2). Hence, the TDA and ML methods are threshold-free. In particular, we investigate the inherent design of TDA methods that provide useful information about topological features of ARs (i.e., AR structure descriptors).

We are also interested in identifying ARs and in understanding how the classification performance of TDA and ML methods varies when these methods are applied to climate model output at different spatial and temporal resolutions. In general, this can make contributions to several climate research problems, for example, evaluating climate model biases in calculating AR statistics under different carbon emission scenarios. Moreover, we compare how outputs of various AR identification methods differ in AR frequency and duration statistics [2, 3]. That is critical to advancing prediction of extreme winter storms

and regional precipitation that has important societal impacts in many areas of the globe.

Note that, in this thesis, we focus on ARs making landfall along the Pacific coast of North America, but the AR recognition method can be extended to other regions on the globe. Our decision to study the Pacific coast of North America was made by the fact that AR phenomena have been extensively investigated on the coast by the climate science community over the last two decades.

To study AR phenomena utilising TDA and ML methods, we address the following research questions (see the full list of research questions in Subsection 1.5.1):

1. Can pattern recognition methods automatically discover regularities in raw climate data to identify AR and AB phenomena? (RQ1)

    a) Do pattern recognition methods based on machine learning models facilitate the characterisation of weather phenomena? (RQ1.a)

    b) To what extent can pattern recognition methods replace weather phenomenon identification methods? (RQ1.b)

    c) Can machine learning models provide the climatological statistics of weather phenomena from global climate model outputs? (RQ1.c)

2. Is it possible to detect ARs by extracting high-level representations of AR phenomena from climate data? (RQ2)

    a) How topological features of ARs can contribute to the characterisation of the structure of ARs? (RQ2.a)

    b) Do machine learning models benefit from topological features to learn high-level representations of ARs from climate model outputs? (RQ2.b)

    c) Are AR pattern recognition methods sensitive to changes in spatiotemporal resolution of climate model simulation outputs? (RQ2.c)

Here, we highlight the main contributions of this chapter and emphasise the novelty compared to the state-of-the-art research:

- We develop an alternative method for AR phenomenon recognition based on TDA and ML methods. To the best of our knowledge, this is the first method based on TDA and ML that has been introduced for recognizing weather patterns in climate datasets.

- We indicate that machine learning models can provide a new way of computing the climatological statistics of AR phenomena. In particular, we analyse the Community Atmosphere Model v5.1 (CAM5.1) and the reanalysis product of the 2nd Modern-Era Retrospective Analysis for Research and Applications (MERRA-2).
- We carry out experiments that show the performance of AR recognition methods is sensitive to changes in spatiotemporal resolution of climate model simulation outputs (i.e. the CAM5.1 outputs).
- We extensively compare our method with other AR identification methods that are a part of the Atmospheric River Tracking Method Intercomparison Project (ARTMIP).

3.1 Atmospheric River Pattern Recognition Method

We propose a supervised method for atmospheric river pattern recognition based on topological data analysis [152, 153] and machine learning [154]. The method consists of two parts: a feature extraction and AR identification, as shown in Figure 3.1. The TDA feature extraction as a first step, followed by an ML classifier to perform the binary classification (i.e., distinguishing AR phenomena from non-AR background). Figure 3.3 shows examples of AR phenomenon and non-AR background. The feature extraction part applies a technique from TDA to extract relevant topological features from two-dimensional climate fields automatically (see Figure 3.2). The TDA technique utilises the Union-Find algorithm [155]. In general, the Union-Find algorithm is applied to different fields of science (e.g., physics, geoscience, and environmental science), because many physical systems, like climate, are modelled on grids or lattices. The algorithm allows to keep track of a set of grid points partitioned into a number of disjoint subsets and perform searching for common elements and joining subsets together [156]. We employ the Union-Find algorithm to extract topological feature descriptors of ARs and non-ARs from climate fields. These topological feature descriptors are called connected regions [157] and are obtained from snapshots of global climate models on a latitude-longitude (regular) grid.

The extracted topological features are provided as an input for the ML classifier. The features of AR phenomena and those that are not ARs are then used in training a Support Vector Machine (SVM) classifier [154, 158]. We note here that the training labels are generated by an heuristic identification method that uses thresholds on Integrated Water Vapour (IWV) fields to identify ARs. The ground truth labels were provided by the Toolkit

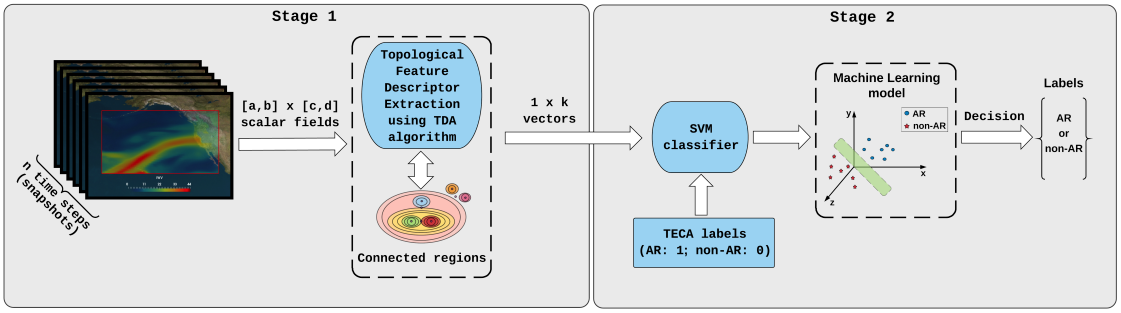


Figure 3.1: The scheme depicts the pattern recognition method for identifying atmospheric river phenomena. The method consists of two steps: topological feature extraction and the machine learning classifier [1].

for Extreme Climate Analysis (TECA) [112]. The toolkit consists of heuristic identification methods for detecting atmospheric phenomena in climate data. For this reason, TECA utilises fixed threshold-based criteria to determine if there is an AR or a non-AR in the given snapshot of climate model output.

The binary classification task is performed using the SVM classifier. The classification task consists of two steps: training the classifier to distinguish ARs from a non-ARs background in snapshots climate fields and testing the constructed SVM classifier on the unlabelled fields to separate ARs and non-ARs, as shown in Figure 3.1.

To assess the performance of the method, we evaluate the SVM classifier performance by means of the following metrics: classification accuracy, precision score (also called positive predictive value) and sensitivity score (also known as recall) .

3.1.1 Topological Feature Extraction

The topological feature extraction part aims to characterize ARs and non-ARs in climate fields automatically, as shown in Figure 3.2. Figure 3.3 shows examples of both AR and non-AR phenomena. Most existing AR identification methods (see Section 2.2) have been designed to use threshold requirements for the identification of ARs. In contrast to those methods, our method proposed here is independent of thresholds by employing topological feature descriptors (i.e., connected regions). We underline that this method is inspired by persistence, which is a concept in TDA that summarizes topological features across all values of a scalar function (or a scalar field) under consideration [152, 153, 159].

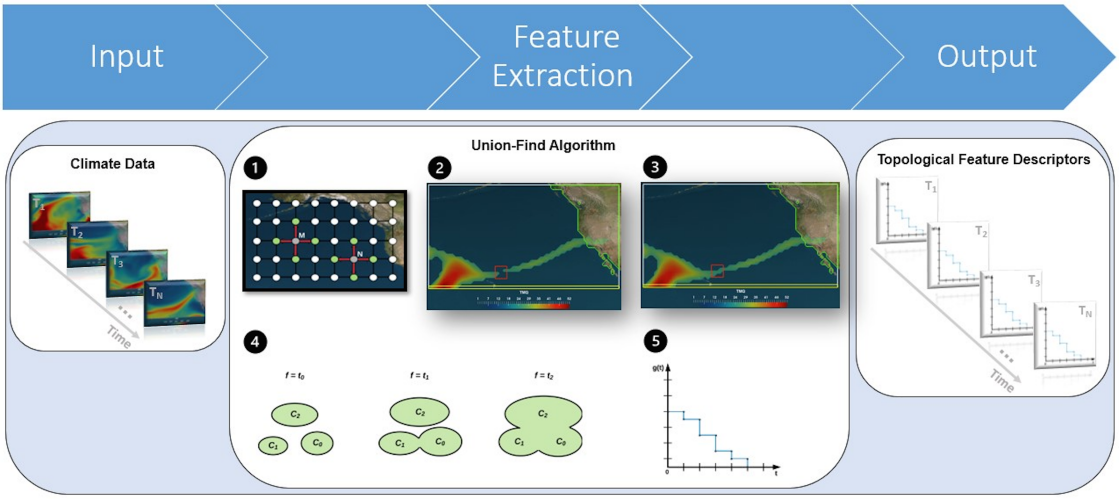


Figure 3.2: The diagram depicts the topological feature extraction from climate data based on the Union Find algorithm. The feature extractor takes input climate fields (i.e., physical variables laid out on the regular grid) at particular time steps $T_1 ... T_N$, where N is the number of time steps (e.g., three hourly). Then, the Union Find algorithm performs a feature extraction from the input data and it outputs topological feature descriptors (i.e., evolution plots).

Climate model data or reanalysis data can be represented as a function from the regular grid to a set of real values (i.e., scalar field), which can be defined, as follows

$$f : [a, b] \times [c, d] \rightarrow [0, I], \tag{3.1}$$

where a, b, c and d are the dimensions of the regular grid. I is the maximal value of physical variable (here $I = 60\ kg\, m^{-2}$), which in our case is Integrated Water Vapour (IWV) over $[0, I]$. Note that the value of I may vary if a different climate field is used, for example, Integrated Water Vapour Transport (IVT) variable.

For gridded climate models or reanalysis data, every grid point has four neighbours in the grid (except boundary points). This is also known as a four-connected neighbourhood, as shown in Figure 3.4. For instance, in terms of point coordinates in the plane, a single grid point at $(x, y) \in [a, b] \times [c, d]$ has four neighbours. Each neighbour has the following coordinates: $(x \pm 1, y)$ or $(x, y \pm 1)$.

Once the mapping of climate data is complete, the feature extractor monitors the evolution of connected regions in a superlevel set at every value t of the function f (here,

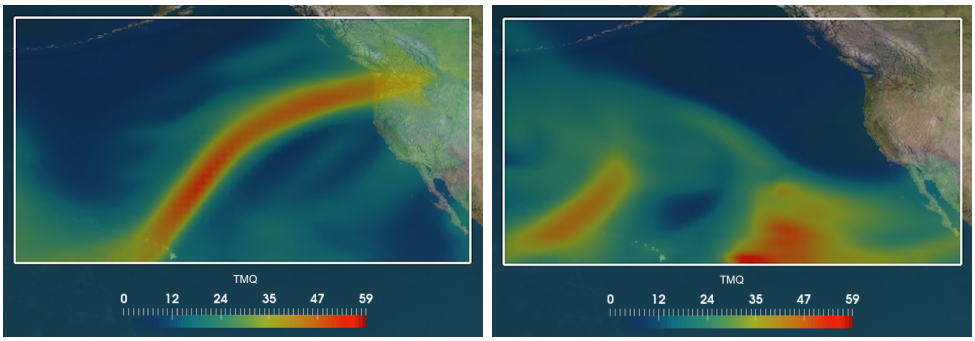


Figure 3.3: Left: An example of atmospheric river (i.e., a long filamentary structures reaching the Pacific coast of North America). Right: An example of non atmospheric river. Shown is the total column integrated water vapour (here, TMQ) in $kg\,m^{-2}$ from a simulation of the Community Atmosphere Model v5.1 at 100 km horizontal resolution [1].

the scalar field). A superlevel set is a set of grid points in the domain of scalar field f with values greater than or equal to t. The superlevel set can be expressed, as follows:

$$f^{-1}[t, +\infty) = \{(x, y) \in [a, b] \times [c, d] : f(x, y) \geq t\}. \tag{3.2}$$

As t value is decreased connected regions of the superlevel set $f^{-1}[t, +\infty)$ appear and grow and eventually merge into larger regions. For example, there are three connected regions (C_0, C_1, C_2) at value t_0 in a superlevel set (defined in Equation (3.2)), as shown in Figure 3.5. As values of f decrease, the component C_0 grows until eventually, at t_1, it merges into the component of C_1, which in turn, merges into the component of C_2 at t_2, and so on.

The above described topological feature extraction of connected regions can be translated into a computational method. It is computed by the Union-Find algorithm [155]. The algorithm finds connected regions by operating on sorted grid point by values of the scalar field f. Here, we sorted the values in decreasing order. The Union-Find algorithm maintains the connected regions and keeps track of the evolution of these regions (i.e., the appearance and disappearance of all connected regions).

We modified the Union-Find algorithm so that it extracts features of ARs and non-ARs from climate fields. The modified algorithm is composed of the following five steps, as shown in Figure 3.2:

❶ Create a new connected region and add the region to the data structure.

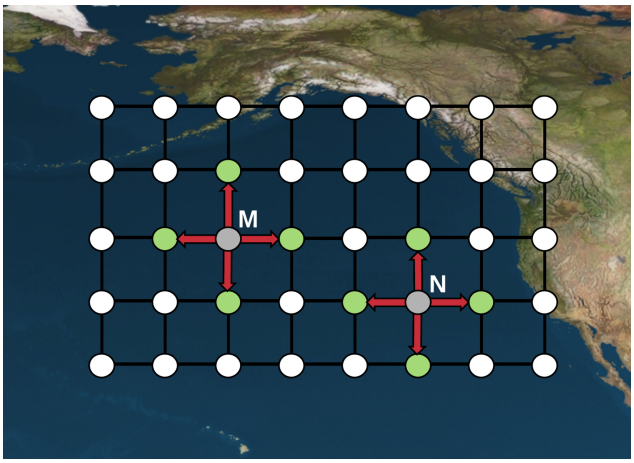


Figure 3.4: An example depicting four-connected neighbourhood that is defined in the latitude-longitude grid in the plane. For instance, each of the grid points M and N (i.e., gray points) has four neighbouring points, i.e. two points along the horizontal direction and two points along vertical direction (i.e., green points) [1].

2. Assign the right connected region to a given grid-point.

3. Check if the connected regions intersect a specified geographical location on the grid. We examine connected regions that intersect the Pacific coast of North America and the latitude of the Hawaiian Islands, as shown on the left side of Figure 3.6.

4. Merge two regions containing at least one identical node into one new connected region, as shown on the right side of Figure 3.6.

5. Keep track of the evolution of a connected region as IWV values are varied (i.e., the number of grid-points that belong to the region) and repeat the four steps mentioned above until all grid-points belong the same connected region.

The computational complexity of the algorithm is $O(n \log(n))$, where n is the number of grid points.

The extracted features of connected regions are encoded in evolution plots. The plots show the recorded number of grid points in growing regions as values of IWV systematically decreases, as shown in Figure 3.7. The horizontal axis t contains values of IWV ($kg\,m^{-2}$). and the vertical axis $g(t)$ shows number of grid points in the connected region. The vectors from the evolution plots are encoded as a matrix with n rows and k columns, where n is

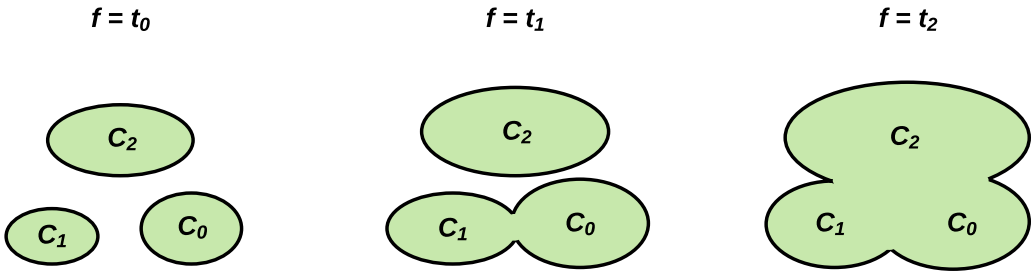


Figure 3.5: An illustration of the connected regions in the superlevel set (defined in Equation (3.2)) that are split into three pieces at value t_0 of f. The regions grow and merge first at value t_1 and then at t_2 when values of scalar field f are gradually decreased [1].

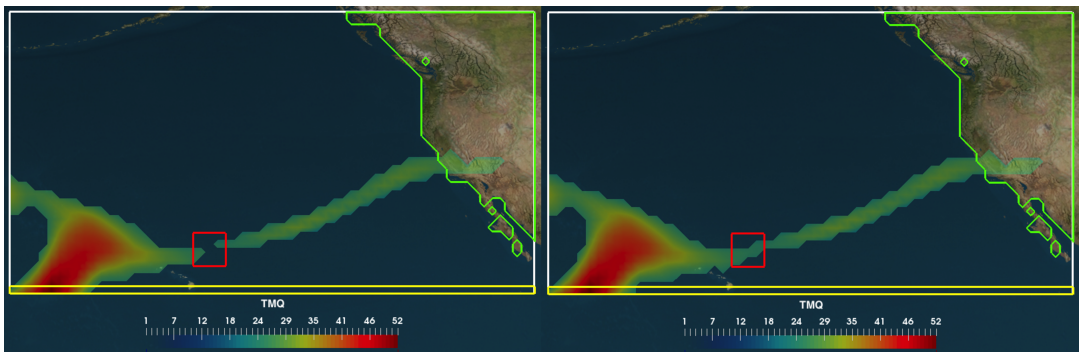


Figure 3.6: An illustration of finding connected regions of an atmospheric river (AR) over a specified geographic location. In this example, the search for ARs is bounded by the latitude of the Hawaiian Islands (here, yellow line) and the Pacific coast of North America (here, green line). Left: The red box indicates location of two regions that are disconnected at some value $f = t_1$. Right: At a new value $f = t_2$, where $t_2 < t_1$, the two connected regions merge into one new connected region forming a valid AR pattern. The IWV ($kg\,m^{-2}$) displayed in this example is from the Community Atmospheric Model v5.1 [1].

the number of time steps and k is the size of the topological feature descriptors returned by the above described algorithm, as is shown in the Figure 3.7. The newly created $n \times k$ matrix is the input data to a ML classifier in the next stage of the AR patter recognition method.

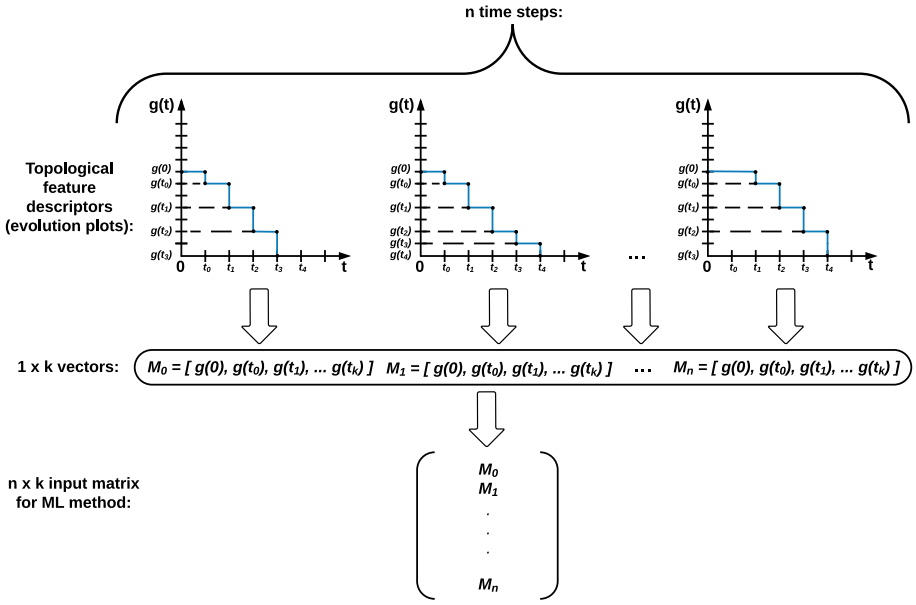


Figure 3.7: Creating an input matrix for a machine learning (ML) classifier. The mapped evolution plots into $1 \times k$ vectors (topological feature descriptors) are stacked on top of each other to construct $n \times k$ input matrix for the ML classifier. Then, the matrix is used as the input by the classifier along with labels provided to an atmospheric river identification method [1].

3.1.2 Supervised Machine Learning Classification

Support Vector Machine (SVM) is one of the most commonly used machine learning methods for pattern recognition tasks [154]. However, deep learning methods have achieved the state-of-the-art performance on pattern recognition tasks in recent years [124], SVM has a couple of advantages when it is applied to a new research problem or dataset. SVM has a smaller number of hyperparameters to tune, an interpretable geometric representation, and produces a unique solution, since the optimization problem is convex [158].

Our main goal is to use SVM as a classifier to perform binary classification. In other words, the SVM decides whether AR phenomena are present or not in given snapshots extracted from global climate fields. The SVM constructs a model based on the labelled topological feature descriptors in the training set and then use it to predict the labels of the descriptors in the testing set. The training part of the SVM aims to find the optimal hyperplane that can separate two groups of patterns (i.e., ARs and non-ARs background). The optimal hyperplane is determined by maximizing the margin between the separating

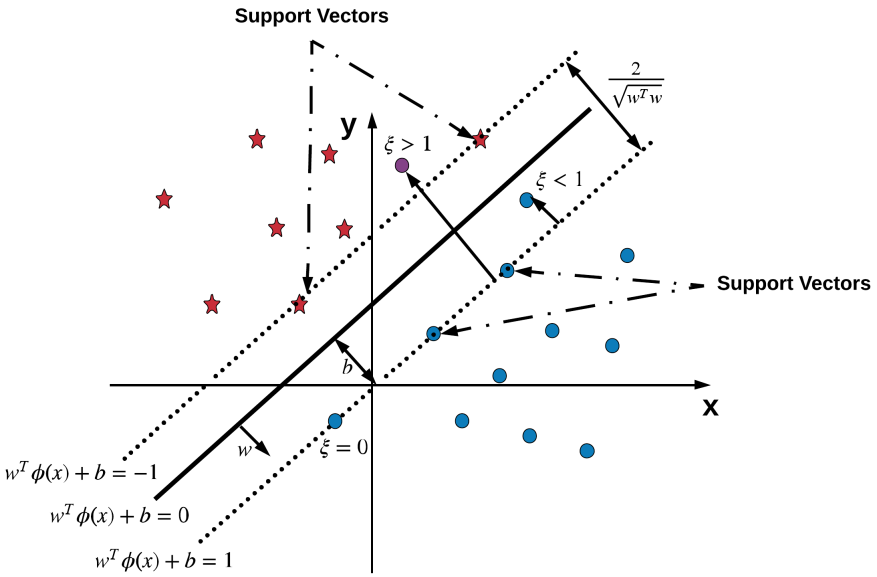


Figure 3.8: An example of linear Support Vector Machine classifier that searches for the optimal hyperplane $w^T\phi(x) + b = 0$, its maximum-margin $\frac{2}{\sqrt{w^T w}}$ by separating samples from two groups in data (here, blue dots and red stars), and all other quantities in the equations (3), (4). ζ is a variable defining how much on the 'wrong' side of the hyperplane a sample is: if it is $1 > \zeta > 0$, the sample is classified correctly, but by less of a margin than the optimal hyperplane, otherwise if it is more than $\zeta > 1$, the sample is classified incorrectly. The magenta dot indicates an example of a misclassified sample from the class of blue dots. Support vectors help to determine the margin for the optimal linear hyperplane. The $\phi(x)$ is a linear transformation in this case [1].

boundary and the training samples closest to it (i.e., support vectors), as shown in Figure 3.8.

Assume a training set of instance-labels pairs (x_i, y_i), $i = 1, ..., N$, where $x_i \in \mathbb{R}^n$ and $y_i \in \{1, 0\}$. The solution of the optimization problem (finding the optimal hyperplane) is given by

$$\min_{w,b,\xi}(\frac{1}{2}w^T w + C\sum_{i=1}^{N}\xi_i), \tag{3.3}$$

subject to

$$y_i(w^T\phi(x_i) + b) \geq 1 - \xi_i \quad \text{and} \quad \xi_i \geq 0. \tag{3.4}$$

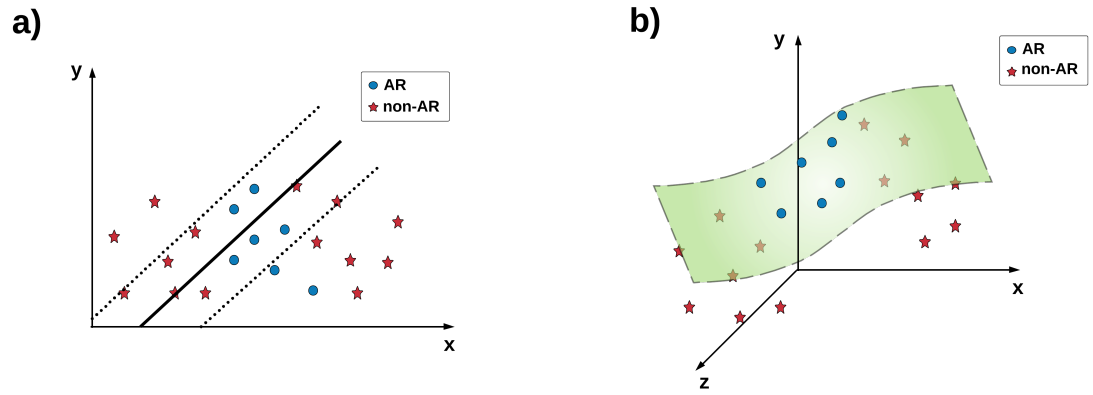


Figure 3.9: a) An illustration of no clear linear separation between two classes (here, ARs and non-ARs) in data. This case cannot be solved using linear SVM. b) Where the set of two groups is not linearly separable in the original space the SVM introduces the notion of a "kernel function induced feature space" which embeds the data into a higher dimensional space where the data are separable [1].

The penalty parameter of the error term takes only values greater than zero ($C > 0$) and $\xi_i \geq 0$ is a minimum error when two groups are not linearly separable (*e.g.*, a noisy signal in training data). The samples $\{x_i\}$, where $x_i \in R^n$, from the training set are mapped into a high dimensional feature space F by means of the transformation $\phi(x_i)$, where $\phi(x) : R^n \rightarrow F$. This transformation makes the samples of two groups (i.e., ARs and non-ARs) separable, as shown in Figure 3.9. Then, the similarity between observations x_i and x_j is computed by kernel function $K(x_i, x_j)$ that can be expressed as an inner product $\langle \phi(x_i), \phi(x_j) \rangle_F$ in the feature space F. Thus, it is sufficient to know $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle_F$ rather than $\phi(x)$ explicitly [160].

In our SVM model, a radial basis function (RBF) kernel is chosen. Because it has been shown that SVM achieves the best results when RBF is the kernel function. The RBF is defined as follows

$$K(x_i, x_j) = exp(-\gamma \left\| x_i - x_j \right\|^2), \quad \gamma > 0, \tag{3.5}$$

where γ is the inverse of the standard deviation of the RBF kernel. The optimal configuration of parameters (C, γ) is found in the experiments by applying loose grid-search and fine grid-search for these two parameters [161].

In this thesis, the matrix created in the feature extraction part of AR identification method is used as the input by the SVM classifier along with the ground-truth labels. The

labels are provided by a heuristic identification method of ARs that is implemented in the Toolkit for Extreme Climate Analysis [112].

3.1.3 Evaluation Metrics

In this subsection, we define the evaluation metrics that were used to assess the reliability of our AR pattern recognition method: classification accuracy score, confusion matrix, precision score, and sensitivity score. We also provide metrics that are used to analyse climatological statistics of ARs, such as frequency and duration of these phenomena. What is more, we describe the preprocessing step of the input to the Support Vector Machine classifier (SVM) to address the issues of imbalanced data [162].

Confusion Matrix

A confusion matrix (also known as a contingency table) is a clear way to present the classification results of ARs concerning testing accuracy of the machine learning classifier. The matrix has two rows and two columns, as shown in Table 3.1. The confusion matrix reports:

	Label non-AR	**Label AR**
Predicted non-AR	True negatives	False positives
Predicted AR	False negatives	True positives

Table 3.1: A confusion matrix (error matrix) is a way to present the performance of the method, especially testing classification accuracy. The matrix reports the number of false positives, false negatives, true positives, and true negatives.

- *False positives* are the cases when the method indicates that an AR is present when it is not present with respect to the ground truth.
- *False negatives* are the cases when the method indicates that an AR is not present, when it is present with respect to the ground truth.
- *True positives* are the cases when the method indicates that an AR is present, when it is present respect to the ground truth.
- *True negatives* are the cases when the method indicates that an AR is not present, when it is not present with respect to the ground truth.

Classification Accuracy

Accuracy is the ratio of correct predictions of ARs to total predictions made by a machine learning classifier. In this thesis, we report *testing accuracy* which is the classification accuracy obtained by applying the trained SVM classifier on testing data. The classification accuracy is defined as follows

$$Accuracy = \frac{True\,positives + True\,negatives}{True\,positives + True\,negatives + False\,positives + False\,negatives} \tag{3.6}$$

Precision Score

Precision is a measure of the classifier repeatability or reproducibility of ARs, and can be computed using a confusion matrix. The score is the ratio of *True positives* to the sum of *True positives* and *False positives*. The precision score is defined as follows

$$Precision = \frac{True\,positives}{True\,positives + False\,positives} \tag{3.7}$$

Sensitivity Score

Sensitivity is a the proportion of actual ARs which are correctly identified as ARs by a classifier. The score is the ratio of *True positives* to the sum of *True positives* and *False negatives*. The sensitivity score is defined as follows

$$Sensitivity = \frac{True\,positives}{True\,positives + False\,negatives} \tag{3.8}$$

Atmospheric River Frequency

Atmospheric river frequency is defined as the percentage of time that a given location experiences the presence of AR phenomena. For instance, if a method produces an AR frequency of 5% at some location, it means that this location experiences the presence of ARs, 5% of the time over a specific period.

Atmospheric River Duration

Atmospheric duration is defined as the continuous length of time that a given location is affected by AR phenomena. For instance, if a method produces an AR duration of 18

hours at this location, it means that the average duration of AR phenomena is 18 hours over a specific period.

Imbalanced Class Data

Balancing the data is motivated by the imbalanced class problem. Each class in the training dataset (here, ARs and non-ARs) is not equally represented in the dataset. This can cause an overfitting problem for the AR recognition method (i.e., ths SVM classifier). In other words, the SVM tends to overfit to the majority class. We circumvent this problem by applying a resampling [163]. The resampling is applied to all matrices created by the feature extraction, along with labeled data generated by the Toolkit for Extreme Climate Analysis (TECA) [112].

3.2 Experimental Results

In this section, we present the results of applying the proposed AR pattern recognition method to climate datasets. The method has been tested on climate model outputs and a reanalysis product. A detailed summary of each dataset is shown in Table 3.2.

Firstly, we analyse the topological feature descriptors of ARs and non-ARs based on the ground truth data. The labels were generated by an objective identification method implemented in the Toolkit for Extreme Climate Analysis (TECA) [112]. Note that the feature descriptors have been normalized to make the comparison of results to different datasets feasible. Secondly, we measure the classification performance and reliability of the Support Vector Machine classifier. To assess the classifier, we calculate the evaluation metrics described in Subsection 3.1.3. Also, we discuss some limitations of the classifier, its typical failure modes, and its precision and sensitivity in identifying ARs. Finally, we provide a comparison of our method to some AR identification methods listed in Table 2.3 and Table 2.4. The results presented here are a part of the Atmospheric River Tracking Method Intercomparison Project (ARTMIP) [2, 3].

Note that, in this thesis, we focus on ARs making landfall along the Pacific coast of North America, but the AR recognition method can be extended to other regions on the globe. Our decision to study the Pacific coast of North America was made by the fact that AR phenomena have been extensively investigated on this coast by the climate science community for over two decades.

3.2.1 Climate Data

In this thesis, we utilise climate model outputs generated by v5.1 of the Community Atmosphere Model (CAM5.1) [164]. The model outputs were produced by the National Energy Research Scientific Computing Center at the Lawrence Berkeley National Laboratory [36]. We also use a reanalysis product, the 2nd Modern-Era Retrospective Analysis for Research & Applications (MERRA-2) data [76]. The reanalysis was provided by the Center for Western Weather and Water Extremes at University of California, San Diego.

The CAM5.1 climate model output is available at three spatial resolutions: 25 km, 100 km, and 200 km. We use both 3-hourly temporal resolution and daily temporal resolution, from January 1980 to June 2017, as shown in Table 3.2. The MERRA-2 reanalysis data are at 50 km spatial resolution and 3-hourly temporal resolution, for the period of January 1980 - June 2017, as shown in Table 3.2.

Climate Dataset	**Time Span**	**Temporal Resolution**	**Spatial Resolution**	**Variable**
CAM5.1 (historical run)	1979-2005	3-hourly and daily	25 km	TMQ
CAM5.1 (historical run)	1979-2005	3-hourly and daily	100 km	TMQ
CAM5.1 (historical run)	1979-2005	3-hourly and daily	200 km	TMQ
MERRA-2 (reanalysis)	1980-2017	3-hourly	50 km	IWV

Table 3.2: Data sources used in the analysis of atmospheric river pattern recognition method: the Community Atmosphere Model v5.1 (CAM5.1) simulations and the 2nd Modern-Era Retrospective Analysis for Research and Applications reanalysis (MERRA-2) dataset. Both TMQ and IWV are a measure of total vertically integrated water vapour in $kg\,m^{-2}$.

In our analysis, we use the total column Integrated Water Vapour (IWV) variable for both the CAM5.1 output and MERRA-2 data. Note that this variable is called the Total Quantity Mass of atmosphere water vapor content (TMQ) for the CAM5.1 output. It can also be called total precipitable water (PRW) in the Climate and Forecast protocols. IWV, TMQ, and PRW are measured in $kg\ m^{-2}$. IWV is derived from the datasets in Table 3.2

and is computed as follows:

$$IWV = -\frac{1}{g}\int_{P_b}^{P_t} q(p)\,dp, \tag{3.9}$$

where q is the specific humidity ($kg\,kg^{-1}$), P_b is 1000 hPa, P_t is 200 hPa, and g is the Earth's mean gravitational acceleration, 9.80665 $m\,s^{-2}$.

Note that this analysis could be performed on other relevant variables, including Integrated water Vapor Transport (IVT). IVT is the vertically integrated vector product of wind and water vapor (i.e., $kg\,m^{-1}\,s^{-1}$) and it is another widely used variable for identifying AR phenomena [127]. However, we note that IWV is observable data by satellite whereas IVT is not.

We choose to use both 3-hourly and daily data because the daily averages often blur certain physical features of AR phenomena. Furthermore, 3-hourly data provides more samples labelled as ARs that is beneficial for training the SVM classifier.

Training the SVM classifier requires labelled data of events that are ARs and those that are not (non-ARs). In other words, each time step (snapshot) has to be tagged with a label one (1) if it contains an AR or zero (0) otherwise. We use the Toolkit for Extreme Climate events Analysis (TECA) [112] to obtain labels for training data. The toolkit uses fixed threshold-based criteria [43] to determine if there is AR in the given snapshot or not. The labels have been generated to for each dataset listed in Table 3.2. It is assumed that labels provided by TECA is ground truth. For each datasets we extracted snapshots of climate fields along the Pacific coast of North America region, as shown in Figure 3.3. Our decision to study this region was made by the fact that AR phenomena have been extensively investigated by other climate science groups working on AR identification methods (see Table 2.3 and Table 2.4).

3.2.2 Classification Results

In this section, we present AR identification results on different datasets which are described in Subsection 3.2.1. We show an analysis of topological feature descriptors of ARs and non-ARs based on the provided ground truth. The feature descriptors are displayed in the form of normalised evolution plots of connected regions to facilitate the analysis. We measure the classification performance and reliability of the SVM classifier. We then provide a comparison of our method to other AR identification methods listed in both Table 2.1 and Table 2.2.

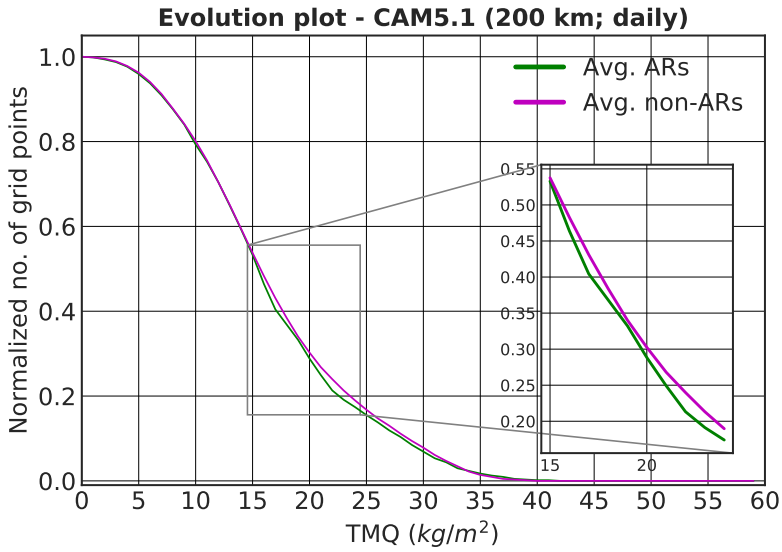


Figure 3.10: An example of normalized evolution plots of averaged topological feature descriptors of TMQ variable (field) of AR and non-AR snapshots. Here, TMQ stands for the Total Quantity Mass of atmosphere water vapor content ($kg\ m^{-2}$) of the Community Atmosphere Model v5.1 dataset (i.e., 200 km spatial resolution and daily temporal resolution) [1].

Figure 3.10 illustrates an evolution plot with two curves of averaged topological feature descriptors. The green and magenta curves correspond to ARs and non-ARs with respect to the ground truth labels. Each curve represents the number of grid points in the connected regions measured by the topological feature extraction. The feature extraction records the evolution of the connected region as a function of TMQ or IWV (see Subsection 3.2.1). We observe that these two curves are close to each other. Hence, it is very challenging to distinguish ARs from non-ARs. However, there is a subtle difference between these two groups of samples from 15 $kg\,m^{-2}$ to 35 $kg\,m^{-2}$ (see the insert region of Figure 3.10), which is a conclusive result with the study of Ralph et al. [43]. That illustrates the discriminative power of topological feature descriptors to identify ARs that can be beneficial in training the SVM classifier. The same analyses using topological feature descriptors has been done for all other datasets listed in Table 3.2.

We now evaluate the performance and reliability of the proposed AR recognition method by measuring the classification accuracy. Table 3.3 and Table 3.4 summarise the

classification accuracy of our method for the CAM5.1 climate model at different temporal resolutions. We note that the SVM classifier has been k-fold-cross-validated (k=10) on each dataset, where eight folds are for training, the one fold is for validation, and the one fold is for testing. In each cross-validation round, different SVM hyperparameters were chosen, and the classifier was trained on a training set, evaluated on a validation set, and tested on a testing set.

Dataset	**Accuracy**	**No of AR**	**No of Non-AR**
CAM5.1 (25 km)	83%	6838	6848
CAM5.1 (100 km)	77%	7182	7581
CAM5.1 (200 km)	90%	3914	3914

Table 3.3: Classification accuracy estimated of Support Vector Machine classifier for the 3-hourly temporal resolution of the Community Atmosphere Model v5.1 with three spatial resolutions: 25 km, 100 km, and 200 km [1].

Table 3.3 shows that the SVM classifier is able to learn to better differentiate ARs from non-ARs when the spatial resolution of the climate model is low. We speculate that the high resolution version of the model more realistically represents AR structures than the low one [36]. That is why IWV fields tend to be noisier, leading to a less smooth topological representation and the decrease in classification accuracy. The high classification accuracy for the CAM5.1 (200 km) suggests that the SVM is able to capture key nonlinear dependencies between topological feature descriptors.

In Table 3.4 we observe a similar trend with the accuracy and model resolution as in Table 3.3. Also, we note that the number of snapshots is about 10 times smaller for the daily temporal resolution datasets than the number of snapshots for the 3-hourly temporal resolution datasets. But this does not affect testing accuracies (consistently above 80%). This suggests that even though event boundaries may be more smeared out in daily averages, the topological feature descriptors sufficiently encode unique information about ARs and non-ARs that SVM is able to distinguish between the two categories with high accuracy. The SVM has the highest classification accuracy for both CAM5.1 (200 km) at the 3-hourly resolution and for CAM5.1 (200 km) at the daily averages.

Table 3.5 shows that the SVM classifier achieved the highest precision and sensitivity scores for 200 km resolution of the CAM5.1 model for both 3-hourly and daily temporal resolutions. The scores are slightly lower for other spatial and temporal resolutions

Dataset	Accuracy	No of AR	No of Non-AR
CAM5.1 (25 km)	82%	624	624
CAM5.1 (100 km)	84%	700	700
CAM5.1 (200 km)	91%	397	397

Table 3.4: Classification accuracy estimated of Support Vector Machine classifier for daily temporal resolution of the Community Atmosphere Model v5.1 with three spatial resolutions: 25 km, 100 km, and 200 km [1].

of CAM5.1 data. The high precision for CAM5.1 with lower resolutions means that the SVM classifier makes fewer mistakes in indicating that ARs exist, when in fact they do not exist (i.e., false positives). For example, Figure 3.12 shows a typical false positive case. This failure mode is often related to the merging of multiple events, either of two ARs or an AR (left panel) and some other event with a high concentration of water vapour and similar topological structure, such as an extra-tropical cyclone (ETC) (right panel). The high sensitivity means that the SVM classifier makes fewer mistakes in misclassifying ARs (i.e, false negatives). For instance, Figure 3.11 shows a typical false negative case. We note that imperfect training data is a challenge in ML, and high-quality ground truth is essential for good model performance. We note that the SVM classifier performs very well in classifying AR phenomena, but has relatively lower performance for non-ARs.

Dataset	Precision	Sensitivity
CAM5.1 (25km, 3-hourly)	0.91	0.74
CAM5.1 (100km, 3-hourly)	0.83	0.67
CAM5.1 (200km, 3-hourly)	0.95	0.85
CAM5.1 (25km, daily)	0.87	0.77
CAM5.1 (100km, daily)	0.86	0.83
CAM5.1 (200km, daily)	0.97	0.85

Table 3.5: Precision and sensitivity scores calculated for all datasets. Both scores show the ability of the Support Vector Machine classifier in assigning correct labels to the test instances [1].

Here, we examine how our method performs in comparison to some AR identification methods, shown in Table 2.1 and Table 2.2. We compare the methods with respect to relevant AR climatologies, such as the frequency and duration of ARs. The results presented here are a part of the Atmospheric River Tracking Method Intercomparison Project

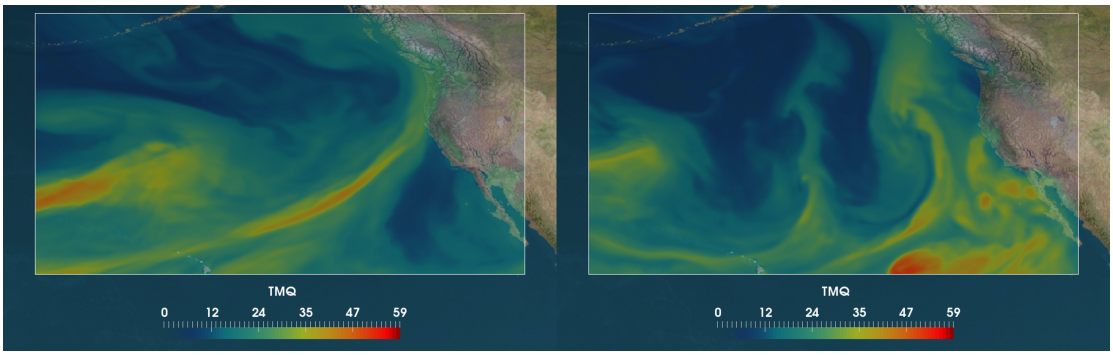


Figure 3.11: Sample images of events from the testing set showing a typical failure mode of the proposed method: examples of atmospheric rivers (ARs) misclassified as non-ARs (false negatives). Figure shows an intergrated water vapour field (here, TMQ) of the Community Atmosphere Model v5.1 (color map) and the land-sea-mass as background (the satellite image). Left: The Support Vector Machine classifier fails likely because there are two separate events in the figure, one is a fully formed AR and another is the start of a new AR; Right: The classifier fails likely due to imperfect training data. Based on the ground truth data, this image is an AR, although (visually) it does not appear to satisfy the definition of an AR. This illustrates how imperfect training data impact the performance of the method because of the limitations of the algorithm used to produce ground truth data [1].

(ARTMIP) [2, 3]. The analysis of all ARTMIP methods was performed on the MERRA-2 reanalysis data. The number of ARs and non-ARs was about 26k instances (i.e., approximately 13k samples per category). The estimated classification accuracy of our AR recognition method is about the same (i.e., 80%) as for datasets of the CAM5.1 model at higher resolutions than 200km. Hence, we demonstrated that the SVM classifier is robust to the source of maps of IWV from CAM5.1 and MERRA-2.

Figure 3.14 presents AR frequency along the Pacific coast of North America and our AR pattern recognition method is noted as “TDA ML”. We observe that the frequency varies greatly as a function of the method used. Furthermore, nearly all methods indicate a rapid increase in AR frequency from a minimum near 32°N to a maximum near 45°N, followed by a more gradual decrease northward toward 56°N. The AR frequency of our method achieves a maximum of just 2% near 39°N and drops down to 0% north of 45°N, where many methods produce their larger frequency values. As we observed before, our method tends to missclassify some ARs (i.e., 0.84 precision score and 0.74 sensitivity score). In the case of the MERRA-2 data the typical failure modes are not different than

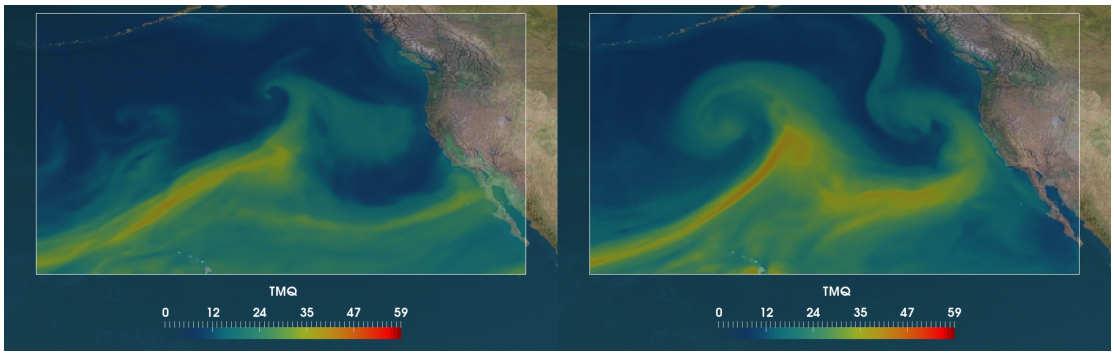


Figure 3.12: Sample images of events from the testing set showing a typical failure mode of the proposed method: examples of non atmospheric rivers (non-ARs) misclassified as ARs (false positives). Figure shows an intergrated water vapour field (here, TMQ) of the Community Atmosphere Model v5.1 (color map) and the land-sea-mass as background (the satellite image). Left: The Support Vector Machine classifier likely fails because of the presence of two AR-like branches that are close to each other, one that has not yet made landfall and another that probably remains after previous event; Right: The classifier fails likely due to the merging of two events, both with high concentration of water vapour, one that appears to be an AR and the other likely an extra-tropical cyclone [1].

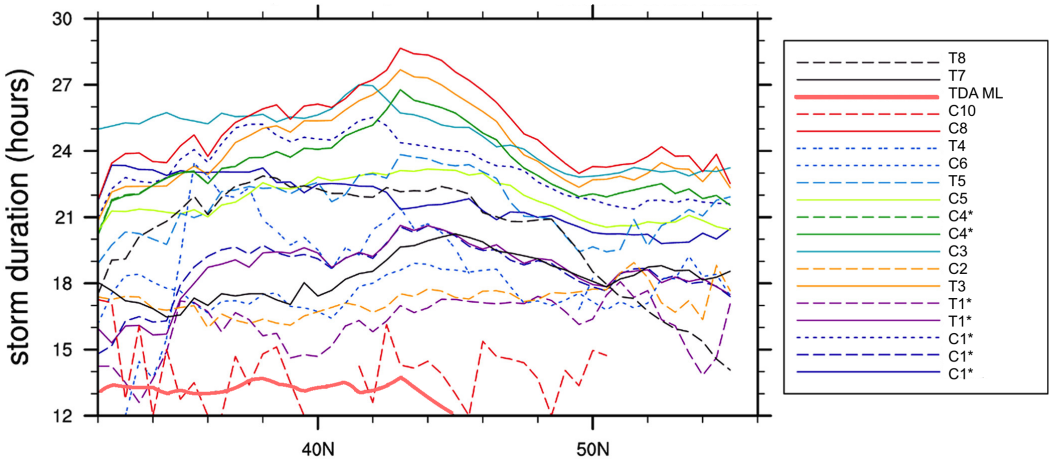


Figure 3.13: AR duration in hours. The results presented here are a part of the Atmospheric River Tracking Method Intercomparison Project (ARTMIP) [2, 3]. Names of methods correspond to method identifiers presented in Table 2.1 and Table 2.2. Our AR pattern recognition method is noted as "TDA ML". The * symbol stands for variations of the same method with the corresponding identifier.

for CAM5.1 datasets.

Our method frequently produces false negatives in recognizing ARs in MERRA-2 data. It is known that the reanalysis data may have higher background noise and low-intensity AR signals. This would result in higher frequency in the more active storm track at latitudes north of 40°N, which may explain the rapid drop-off in AR detection associated with this method. It is also worth mentioning that our method excels at lower latitudes, where it identifies a relatively larger number of ARs than most AR identification methods.

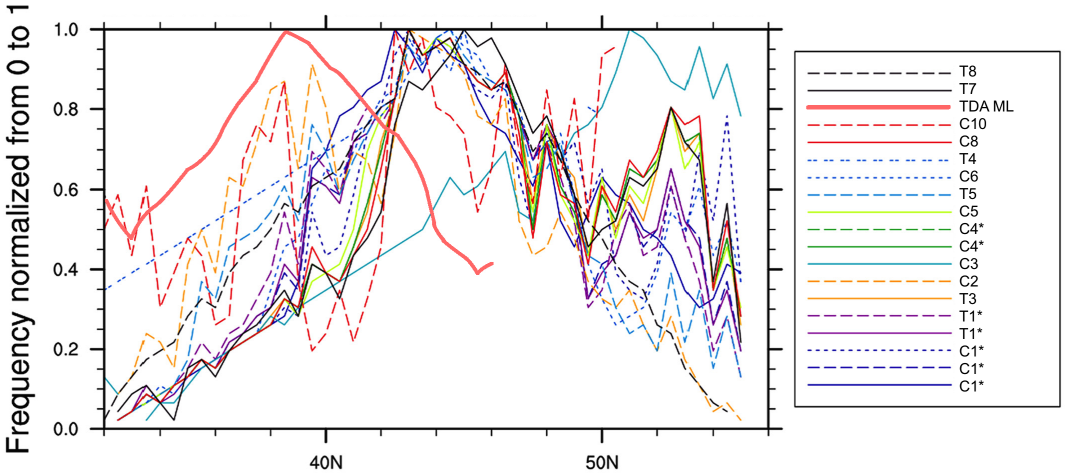


Figure 3.14: Percentage of period along AR transect. The results presented here are a part of the Atmospheric River Tracking Method Intercomparison Project (ARTMIP) [2, 3]. Names of methods correspond to method identifiers presented in Table 2.1 and Table 2.2. Our AR pattern recognition method is noted as "TDA ML". The * symbol stands for variations of the same method with the corresponding identifier.

Figure 3.13 shows AR duration along the Pacific coast of North America. Most methods exhibit a gradual increase in AR duration from 32°N to a maximum near 42–44°N, followed by a gradual decline northward toward 50°N. This trend in the duration matches the maximum in AR frequency characteristic.

As we noted before, our method typically identifies fewer ARs than other methods used in the intercomparison. Furthermore, the identified ARs have shorter average durations than for other AR identification methods.

3.3 Discussion and Summary

In this chapter, we proposed a novel and automated method for identifying AR phenomena in climate datasets. The method combines topological feature extractor with SVM classifier. We demonstrated that the proposed method is reliable, robust and performs well by testing it on a wide range of spatial and temporal resolutions of the CAM5.1 climate model output as well as the MERRA-2 reanalysis product. The performance of the method is measured by classification accuracy, precision and sensitivity scores. TECA provided the ground truth labels for training the SVM classifier. We observed that the SVM classifier is biased by the ground truth data produced by TECA using the threshold-based criteria for identifying of ARs. We also note that characterizing the influence of using different ground truth data is beyond the scope of this thesis.

Despite background noise, low-intensity AR signals and the existence of other phenomena within the two-dimensional snapshots of climate fields, our method achieved high accuracy. The method achieved the highest accuracy for low-resolution climate data. We speculate that this is because high-resolution simulations produce noisy spatial patterns that mislead the machine learning model more easily than low-resolution simulations.

The key advantage of the topological feature descriptors used in this work is a threshold-free method that succinctly encapsulates the most important topological features of ARs. We note that the method does not have a threshold requirement for the TDA step. Moreover, when the spatial resolution of the climate model changes, there is no need for threshold parameter retuning, unlike in the case of engineered heuristic methods used by most AR identification methods.

Chapter 4

Atmospheric Blocking Phenomenon Identification in Climate Reanalysis Products

In this chapter, we focus on deep learning approaches for identifying Atmospheric Blocking (AB) phenomena in climate reanalysis data [4]. In particular, we are interested in exploring different architectures of deep Convolutional Neural Networks (CNNs) for detection and localisation of ABs. CNNs are capable of automatically extracting high-level representations of ABs from climate data to characterise, recognise, and localise AB phenomena. Therefore, CNNs overcome key critical challenges in designing objective identification methods of ABs, such as the selection of appropriate climate fields and translating complex Rex's criteria into detection schemes (see Chapter 2.3). Here, we utilise deep CNNs to detect and localise AB phenomena in climate reanalysis products. We are also interested in observing how the performance of CNNs varies between hemispheres and various regions over the globe. In general, this research may help in evaluating AB phenomena climatology regarding the geographical distribution of the frequency of occurrence of ABs.

Moreover, we investigate a relationship between the inherent design of deep CNN architectures, which is the network depth, and CNN performance in identifying AB phenomena. For this reason, we perform extensive experiments to evaluate which CNN architectures perform the best. In other words, we want to observe a statically significant increase in AB classification accuracy and a statically significant decrease in prediction error of local-

isation parameters of ABs. This study may serve as a guide in the design and selection of CNN architectures for AB pattern recognition methods.

To investigate AB phenomena and to formulate a pattern recognition task for identifying ABs, we address the following research questions (see the full list of research questions in Subsection 1.5.1):

1. Can pattern recognition methods automatically discover regularities in raw climate data to identify AR and AB phenomena? (RQ1)

    a) Do pattern recognition methods based on machine learning models facilitate the characterisation of weather phenomena? (RQ1.a)

    b) To what extent can pattern recognition methods replace weather phenomenon identification methods? (RQ1.b)

    c) Can machine learning models provide the climatological statistics of weather phenomena from global climate model outputs? (RQ1.c)

3. Is it possible to detect ABs by learning high-level representations of AB phenomena from multivariate and high-dimensional climate reanalysis data? (RQ3)

    a) Can deep learning models automatically extract high-level representations of ABs to detect these phenomena in climate reanalysis products? (RQ3.a)

    b) Is it possible to predict the localisation and geometric properties of ABs in climate reanalysis data? (RQ3.b)

    c) To what extent can the depth of deep learning architectures contribute to the AB detection performance in various geographical regions over the globe? (RQ3.c)

Here, we highlight the main contributions of this chapter and emphasise the novelty compared to the state-of-the-art research:

- We develop a new pattern recognition method for the detection and localisation of AB phenomena. The method is based on the state-of-the-art deep CNN architectures. To the best of our knowledge, this is the first method based on CNNs that has been introduced for identifying ABs in different regions over the globe.

- We investigate five different architectures of the generic CNN-based classifier and CNN-based regressor respectively designed for each stage of the AB recognition

method. In particular, we apply the classifier and the regressor to the European Centre for Medium-Range Weather Forecasts Interim Reanalysis (ERA-Interim) product.

- We show that the AB detection performance significantly increases for the deep CNN architectures. In contrast to that, we see that the estimation error of AB location significantly decreases in the localisation problem for the shallow CNN architectures.
- We conduct experiments and demonstrate that the proposed CNN architectures tend to achieve better AB detection and localization performances in regions of the Northern Hemisphere than in regions of the Southern Hemisphere.

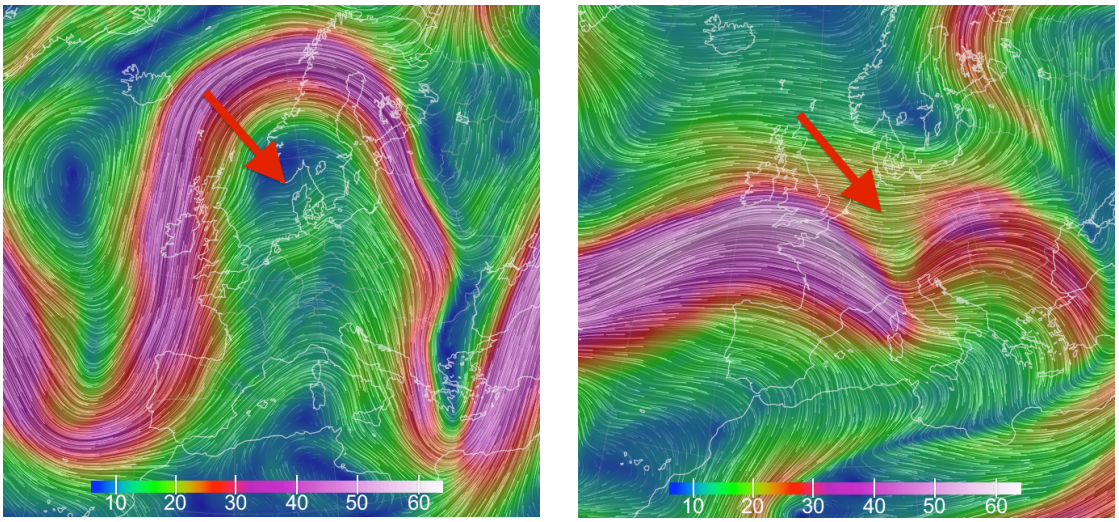
4.1 Atmospheric Blocking Pattern Recognition Method

In this section, we propose a hierarchical pattern recognition method for identifying atmospheric blocking (AB) phenomena [4]. Figure 4.1 shows an example of AB phenomenon and non-AB situation. The method is based on deep Convolutional Neural Network (CNN) architectures [101], inspired by Visual Geometry Group (VGG) architectures [165].

The VGG architectures are based on an analysis of the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting. The information flow in a typical VGG architecture is that input data are processed by several modules of convolutional layers and pooling layers that extract feature representations. The feature representations are then fed into one or more fully connected layers that make classification decisions or predictions. VGGs are the most commonly referred CNN architectures in the deep learning literature, and one of the state-of-the-art CNN architectures in classification and localisation tasks.

The proposed AB pattern recognition method consists of two stages: AB detection defined as a binary classification task and localisation of ABs formulated as a regression problem, as shown in Figure 4.2. The AB detection, as a first part, uses a CNN-based classifier that distinguishes ABs from non-ABs in different regions over the globe. Samples with detected AB phenomena are passed to the second stage. In the second stage, a CNN-based regressor predicts AB location parameters in the climate fields: a mass centre (a latitudinal position and a longitudinal position) and a radius of a minimum enclosing circular box of AB in various regions on the globe.

We note here that the training labels for the CNN-based classifier are provided by an objective method of Schwierz et al. [149] that uses thresholds on potential vorticity and



(a) (b)

Figure 4.1: Right: An example of atmospheric blocking marked by the red arrow (i.e., a high pressure pattern that blocks and diverts the jet stream). Here, it is shown a particular example of the Omega block that forms a shape resembling the Greek letter omega; Left: An example non atmospheric blocking phenomenon that does not cause unusual changes in the normal flow of the westerly winds (marked by the red arrow); Shown is the wind speed field ($m\,s^{-1}$) at $250\,hPa$ pressure level (i.e., the elevation) from a simulation of the Global Forecast System produced by the National Centers for Environmental Prediction. Source: The US National Weather Service data visualised by `http://earth.nullschool.net`.

space-time criteria to identify ABs. The classification task consists of two steps: training and validation of the CNN-based classifier to distinguish ABs from non-ABs in snapshots of climate fields, and testing the constructed classifier on the unlabelled fields to indicate ABs and non-ABs. Furthermore, the CNN-based regressor predicts location parameters of AB (i.e., a mass centre and a radius of a minimum enclosing circular box) passed from the CNN-based classifier. We note that the regression task also consists of training, validation, and testing. The location parameters of minimum enclosing circular boxes were provided by a computational geometry algorithm [166] applied to the output of the objective method of Schwierz et al. [149].

To assess the performance of the method, we evaluate the CNN-based classifier and CNN-based regressor performances by means of the following metrics: classification accuracy and F1 score (also known as the harmonic mean of precision and sensitivity) for

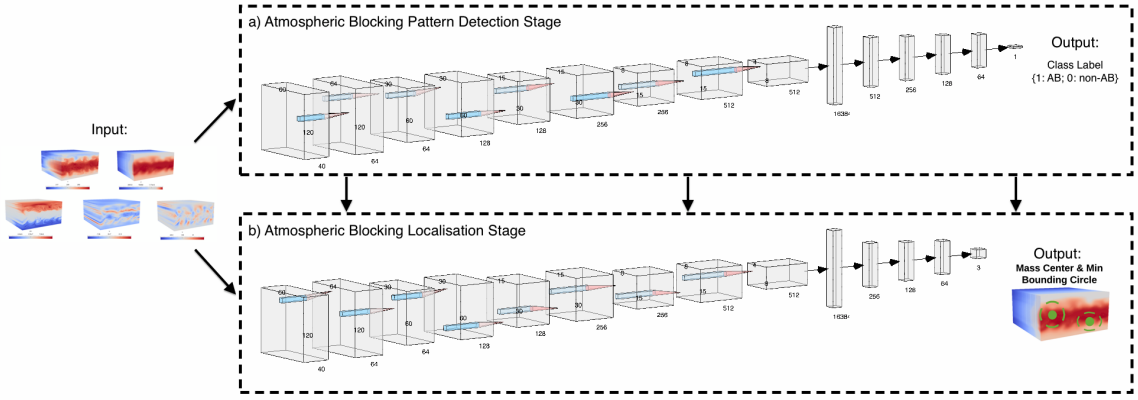


Figure 4.2: The scheme depicts the hierarchical atmospheric blocking (AB) pattern recognition method. The method consists of two parts: (a) a convolutional neural network (CNN) based classifier that distinguishes ABs and non-ABs; and (b) a CNN-based regressor that predicts three parameters describing AB location, i.e. a latitudinal position, a longitudinal position, and a radius of a minimum enclosing circular box [4].

the classifier, and concordance correlation coefficient and mean percentage error for the regressor. We also employ statistical testing to observe a statistically significant increase or decrease in CNN architectures performance for different CNN depths and various regions on the globe.

4.1.1 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are one of the most commonly used deep learning models that have achieved recent state-of-the-art results on visual pattern recognition tasks [101, 124]. The novelty that was introduced by CNNs is that the dual-step approach to the classification problem has been replaced. The dual-step approach relies on extracting handcrafted feature descriptors from images that serve as an input to a machine learning classifier. In contrast to this traditional approach, CNN models utilise multiple stages of automatic information processing for feature extraction, transformation, and for pattern analysis, for example, image classification [167, 168].

In general, CNNs have a topology of feedforward neural networks that process data in one direction only, i.e. from their inputs to their outputs. They consist of convolutional layers and subsampling layers (e.g., maximum pooling layers), which are grouped into modules followed by one or more fully connected layers. When several modules are stacked

on top of each other, they form a deep CNN architecture, for example, Visual Geometry Group (VGG) architectures [165]. The information flow in a typical CNN architecture is that input data are processed by several modules of convolutional layers and pooling layers that extract feature representations. The feature representations are then fed into one or more fully connected layers that make classification decisions or predictions.

Convolutional layers are the main feature extraction components of CNN architectures. They extract the high-level feature representations of their input data, here, images or scalar fields. The layers consist of nodes (also sometimes called "neurons"), which outputs are arranged into feature maps. Each node in a feature map has a receptive field, which is connected to a neighbourhood of nodes in the previous layer via a set of convolutional filters (also called weights). When input images are convolved with the filters, a new feature map is produced, and the results of the convolution operation are passed through a nonlinear activation function. Note that different feature maps within the same convolutional layer can have different filters so that several features can be extracted.

In other words, each layer is a place where two main operations are performed. First, convolutional layers compute a linear combination of the outputs of the previous layers or input data if it is the first layer. Then the linear combination is transformed by non-linear activation functions. Formally, the output feature maps can be formulated as follows:

$$\mathbf{Y} = h(\mathbf{W} * \mathbf{X}), \tag{4.1}$$

where the input image is denoted by $\mathbf{X}$ and the convolutional filter related to a given feature map is denoted by $\mathbf{W}$. The multiplication sign $(*)$ refers to the two-dimensional convolutional operator. The operator is used to calculate the inner product of the filter at each location of the input image. $h(\cdot)$ represents the nonlinear activation function.

CNNs are a succession of linear combinations followed by nonlinear activation functions. These functions allow for the extraction of nonlinear features of inputs. Here, we provide the activation functions that are used in this chapter: a rectified linear unit (RELU) function, a sigmoid function, and a hyperbolic tangent function. RELU activation function is defined, as follows:

$$h(x) := \begin{cases} 0 \text{ for } x < 0 \\ x \text{ for } x \geq 0 \end{cases}, \tag{4.2}$$

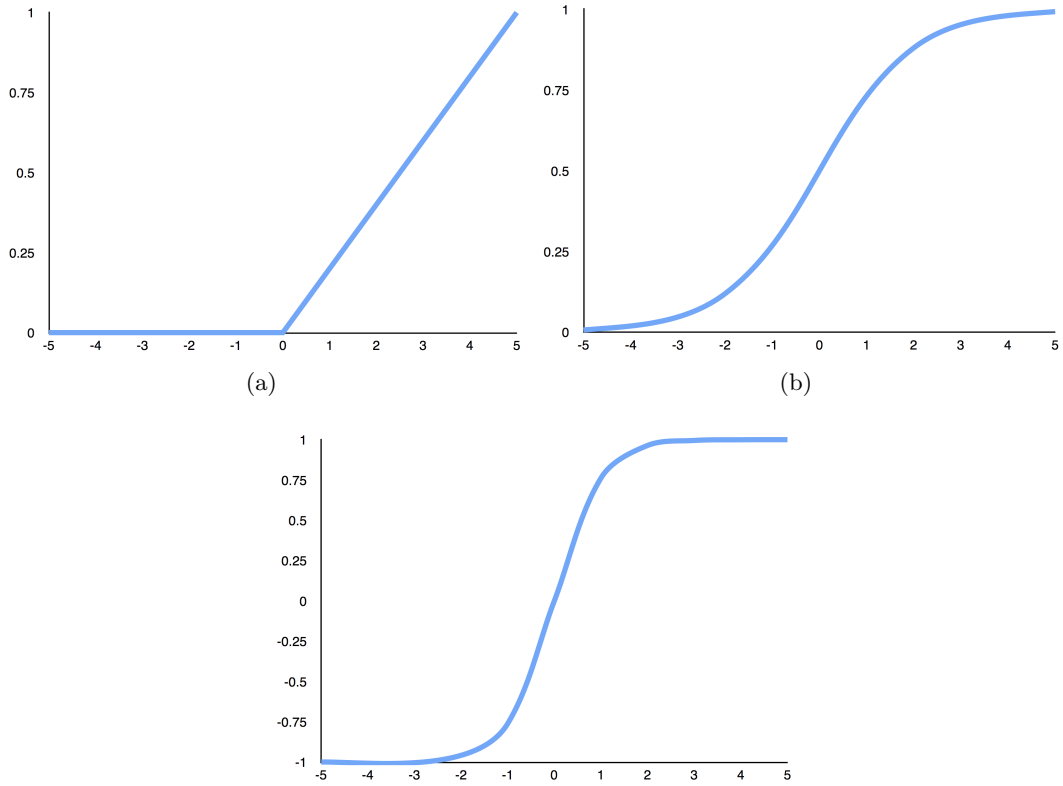


Figure 4.3: Activation functions: (a) a rectified linear unit (RELU); (b) a sigmoid function (the logistic function); (c) the hyperbolic tangent.

where x is the element of convolved image (see Figure 4.3 (a)). The sigmoid activation function is given by:

$$h(x) := \frac{1}{1 + e^{-x}}, \tag{4.3}$$

where x is the element of convolved image, and h returns values in the range $(0, 1)$ (see Figure 4.3 (b)). The hyperbolic tangent activation function is defined in the following way:

$$h(x) := tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}, \tag{4.4}$$

where x is the element of convolved image, and h returns values in the range $(-1, 1)$ (see Figure 4.3 (c)).

After convolutional layers, pooling layers are the most important components of CNN architectures. The main purpose of the pooling layers is that they reduce the spatial dimensions of the feature maps. Hence, the layers help achieve spatial invariance of CNNs to input distortions and translations. Max pooling layers are the most widely used in the deep learning literature [169, 170]. They propagate the maximum value within a receptive field to the next layer in CNN architectures. Formally, the max pooling operation selects the largest element within each receptive field, as follows:

$$\mathbf{Z}^{out}(i,j) = \max_{(p,q)\ \in\ \Omega_{i,j}} \mathbf{Z}^{in}(p,q), \tag{4.5}$$

where $\mathbf{Z}^{out}$ is the output of the pooling operation associated with the feature map. $\mathbf{Z}^{in}$ denotes the element of the input feature map at location (p,q) contained by the pooling region $\Omega_{i,j}$, which surrounds a receptive field around the position (i,j).

The fully connected layers that follow convolutional and pooling layers interpret abstract feature representations by performing the function of high-level reasoning in CNN architectures.

In general, training in CNNs is achieved by finding the parameters that minimise a specific loss function (also called objective or cost function). For this purpose, learning algorithms are used to adjust their free parameters (i.e., the weights of filters) in order to attain the desired CNN output.

In this chapter, we use two most commonly used loss functions, i.e. a binary cross-entropy function for binary classification and a mean square error function for regression. Here, we provide mathematical expression for these loss functions.

Binary Cross-Entropy (BCE) is a measure of the dissimilarity between the true label probability and the predicted label probability. BCE is given by:

$$L_{BCE}(y,y') := -\sum_{i=1}^{N} y\ log(y'), \tag{4.6}$$

where y is true label probability, y' is predicted label probability, and N is the number of observations.

Mean Squared Error (MSE) is the sum of the square of the differences between the predicted value and the true value divided by the number of observations. MSE loss is

defined as follows:

$$L_{MSE}(y, y') := \frac{1}{N}\sum_{i=1}^{N}(y - y')^2, \tag{4.7}$$

where y is the true value, y' is the predicted value, and N is the number of observations.

We minimise the loss functions by a backpropagation algorithm, for example, the stochastic gradient descent algorithm [171]. The algorithm adjust a network's parameters in order to minimize errors that affect CNNs performance.

Convolutional Neural Network Classifier

CNN Architectures				
A	B	C	D	E
Input: (60 × 120 × 40)				
conv-64	conv-64 conv-64	conv-64 conv-64	conv-64 conv-64	conv-64 conv-64
Maxpooling				
conv-128	conv-128	conv-128 conv-128	conv-128 conv-128	conv-128 conv-128
Maxpooling				
conv-256	conv-256	conv-256	conv-256 conv-256	conv-256 conv-256
Maxpooling				
conv-512	conv-512	conv-512	conv-512	conv-512 conv-512
Maxpooling				
FC-512				
FC-256				
FC-128				
FC-64				
Output: FC-1				

Table 4.1: Convolutional Neural Network (CNN) architectures are shown in columns. The depth of the architectures increases from the left (A) to the right (E). The parameters of both types of layers are denoted as follows: conv-(number of filters); FC-(number of channels) [4].

In the first stage of the AB pattern recognition method, we develop a generic CNN-based classifier, which we refer to as architecture A, as outlined in Table 4.1. The architecture has a total of eight weight layers, including four convolutional (conv.) layers and four Fully Connected (FC) layers.

Each convolutional layer is followed by a max pool layer (Maxpooling). The width of convolutional layers starts from 64 filters in the first layer and then increases by a factor of two after each Maxpooling layer, until it reaches 512 filters. In contrast, the width of FC layers starts from 512 channels and then decreases by a factor of two, until it reaches 64 channels. The output FC layer performs a binary classification and therefore contains one value, i.e. an AB label or a non-AB label.

Four architectures of a CNN-based classifier follow the architecture A and gradually increase the number of convolutional layers (their depth). The configuration of the FC layers is the same in all architectures. All architectures are referred by their names (A-E) and all details on the architectures are listed in Table 4.1.

Convolutional Neural Network Regressor

In the second stage of the AB pattern recognition method, we develop a generic CNN-based regressor, which we refer to as architecture A, as outlined in Table 4.2. The architecture has a total seven weight layers, including four convolutional (conv.) layers and three Fully Connected (FC) layers.

Each convolutional layer is followed by a max pool layer (Maxpooling). The width of convolutional layers starts from 64 filters in the first layer and then increases by a factor of two after each Maxpooling layer, until it reaches 512 filters. In contrast, the width of FC layers starts from 256 channels and then decreases by a factor of two, until it reaches 64 channels. The output FC layer performs a multi-regression task and therefore contains three parameter values:

- a latitudinal position - the position north or south of the equator;
- a longitudinal position - the position east or west of the prime meridian;
- a radius of a minimum enclosing circular box - the smallest radius of the circle that contains a given AB phenomenon.

Four architectures of a CNN-based regressor follow the architecture A and gradually increase the number of convolutional layers (their depth). The configuration of the FC layers

CNN Architectures				
A	B	C	D	E
Input: (60 × 120 × 40)				
conv-64	conv-64 conv-64	conv-64 conv-64	conv-64 conv-64	conv-64 conv-64
Maxpooling				
conv-128	conv-128	conv-128 conv-128	conv-128 conv-128	conv-128 conv-128
Maxpooling				
conv-256	conv-256	conv-256	conv-256 conv-256	conv-256 conv-256
Maxpooling				
conv-512	conv-512	conv-512	conv-512	conv-512 conv-512
Maxpooling				
FC-256				
FC-128				
FC-64				
Output: FC-3				

Table 4.2: Convolutional Neural Network (CNN) architectures are shown in columns. The depth of the architectures increases from the left (A) to the right (E). The parameters of both types of layers are denoted as follows: conv-(number of filters); FC-(number of channels) [4].

is the same in all architectures. All architectures are referred by their names (A-E) and all details on the architectures are listed in Table 4.2.

4.1.2 Evaluation Metrics

To assess the performance of architectures of a CNN-based classifier and a CNN-based regressor, we use the following metrics: a classification accuracy and F1 score for a classification task, and concordance correlation coefficient and the mean percentage error for a regression problem. Note that we defined a confusion matrix needed for calculating classification accuracy and F1 score in Chapter 3: Section 3.1.3.

F1 Score

The F1 score measures a weighted harmonic mean of the precision and sensitivity scores of a classifier. Note that these scores were defined in Chapter 3: Section 3.1.3. Here, F1 score is defined as follows

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{4.8}$$

Concordance Correlation Coefficient

The concordance correlation coefficient (CCC) is the measure of agreement between the true variables and the predicted variables by a regressor. CCC is defined as follows

$$CCC_{m,n} = \frac{2\rho\sigma_m\sigma_n}{\sigma_m^2 + \sigma_n^2 + (\mu_m - \mu_n)^2}, \tag{4.9}$$

where μ_m and μ_n are the means for two variables m and n, and σ_m and σ_n are the corresponding variances. ρ is the Pearson correlation coefficient between the two variables.

Mean Percentage Error

The mean percentage error (MPE) is a measures error of a regressor if it systematically underestimates or overestimates the predicted variables. The error is defined as follows

$$MPE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - y_i'}{y_i}\right) \times 100\%, \tag{4.10}$$

where y is the true value and y' is the predicated value of a given quantity, and N is the number of observations.

Statistical Significance Testing

We use McNemar's statistical test at a significance level of 0.05 on the output of classification task [172] and Wilcoxon Signed-Rank test at a significance level of 0.05 for the output regression task [173]. Both tests evaluate which CNN architecture performs the best, i.e. we observe a statically significant increase and decrease in AB classification accuracy and error of predicting localization parameters of ABs, respectively. To facilitate the comparison, CNN architectures have been ranked according to accuracy values for the classification task and the respective MPE values for the regression task.

4.2 Experimental Results

In this section, we present the obtained results for detecting and localising AB phenomena in climate reanalysis products. A detailed summary of the reanalysis data is shown in Table 4.3. To identify AB phenomena, we apply a hierarchical pattern recognition method that consists of two stages: the detection stage that is formulated as a binary classification and the localisation stage that is defined as a regression task. We explore five CNN architectures and evaluate their performance in six geographical regions over the globe.

Firstly, we analyse the results of a CNN-based classification task in distinguishing ARs and non-ARs. Labels for that task were generated by an objective identification method of Schwierz et al. [149]. We also measure the classifier performance and reliability by calculating evaluation metrics, such as classification accuracy and F1 score, described in Subsection 4.1.2.

Secondly, we analyse the results of a CNN-based regression task in predicting ARs location parameters: a latitude, a longitude and a radius of a minimum bounding circular box. Samples with identified AB phenomena were passed from the CNN-based classifier. Furthermore, we measure the regressor performance by calculating evaluation metrics, such as CCC and MPE, described in Subsection 4.1.2.

Note that we provide the analysis of five different CNN architectures for both classification task and regression task. We evaluate their performance in six geographical regions over the globe (see Figure 4.5). We also note that the CNN-based classifiers and regressors have been k-fold-cross-validated (k=10) on the dataset, where eight folds are for training (80%), the one fold is for validation (10%), and the one fold is for testing (10%). In each cross-validation round, different CNN hyperparameters were chosen, and the classifiers and regressors were trained on a training set, evaluated on a validation set, and tested on a testing set.

4.2.1 Climate Data

In our analysis, we use the European Centre for Medium-Range Weather Forecasts Interim Reanalysis (ERA-Interim) products. We use five physical variables on a regular grid described in Table 4.3, each at eight different pressure levels in millibars (mb): 150 mb, 200 mb, 250 mb, 300 mb, 350 mb, 400 mb, 450 mb, and 500 mb, as shown in Figure 4.4. The variables are at six-hourly timesteps at approximately 80 km spatial resolution (180 pixels $\times$ 360 pixels based on T119 spectral model grid resolution) from January 1, 1980

to December 31, 2016. The dataset of identified ABs is based on an identification method

Variable	Variable units	Description
T	K	Air temperature
Z	$m^2\ s^{-2}$	Geopotential
PV	$K\ m^2\ kg^{-1}\ s^{-1}$	Potential vorticity
U	$m\ s^{-1}$	Zonal wind
V	$m\ s^{-1}$	Meridional wind

Table 4.3: Available variables of ERA-interim reanalysis products.

by Schwierz, et al. [149]. It was provided by the Institute for Atmospheric and Climate Science at the ETH Zurich, Switzerland. The method consists a semi-automatic detection method that outputs 2D blocking indices and was applied to the ERA-Interim data. The output of this procedure yields 2D binary fields or masks with the value one (1) at pixels that meet the method criteria for ABs and the value zero (0) at pixels that do not (i.e., non-ABs).

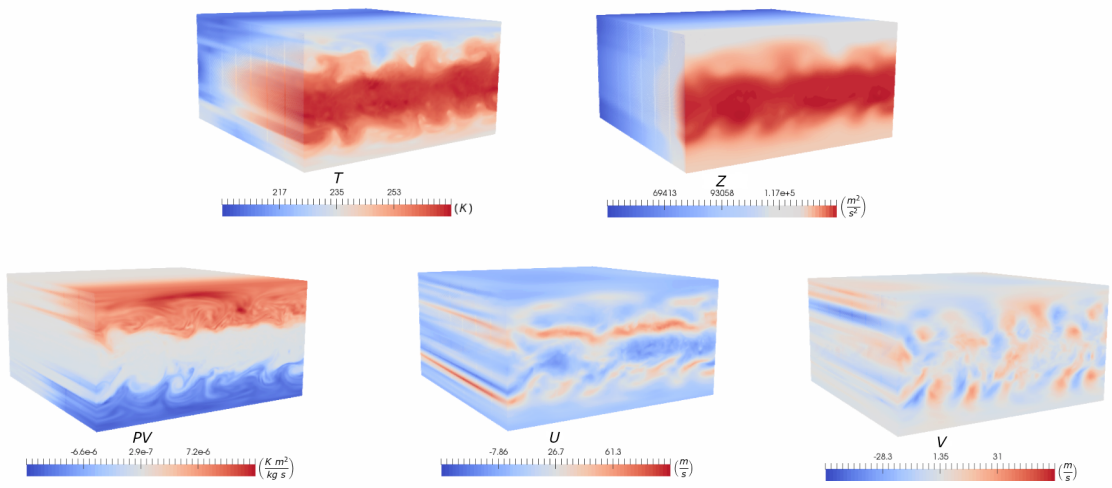


Figure 4.4: A visualisation of available ERA-interim variables: T - the air temperature, Z - the geopotential, PV - the potential vorticity, U - the zonal winds, and V - the meridional winds.

We generated the ground-truth labels (i.e., ABs and non-ABs) based on the output of the above mentioned procedure. The global image of 180 pixels × 360 pixels × 40 channels is divided into six images (regions) of size 60 pixels × 120 pixels × 40 channels.

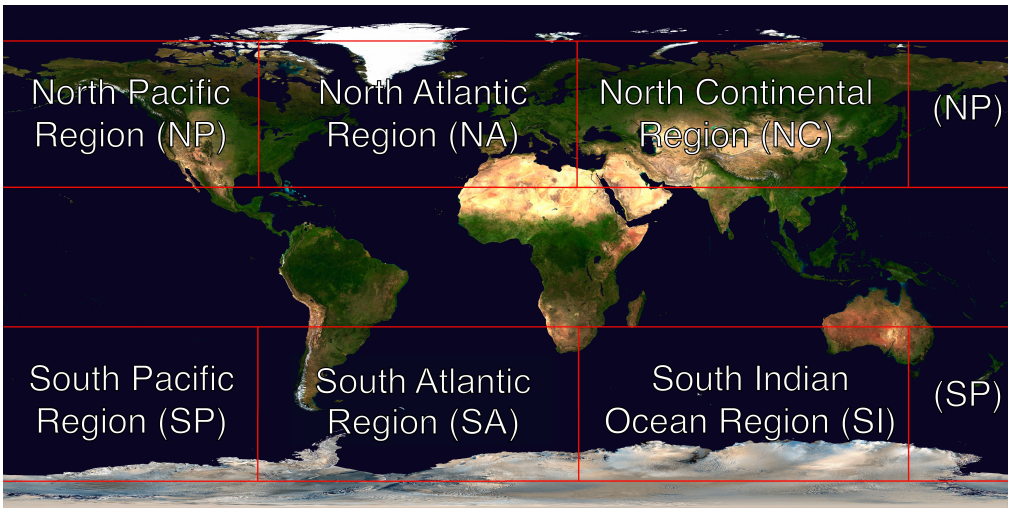


Figure 4.5: An example of the world map (globe) with six defined regions of interest: the North Pacific Region (NP); the North Atlantic Region (NA); the North Continental Region (NC); the South Pacific Region (SP); the South Atlantic Region (SA), and the South Indian Ocean Region (SI) [4].

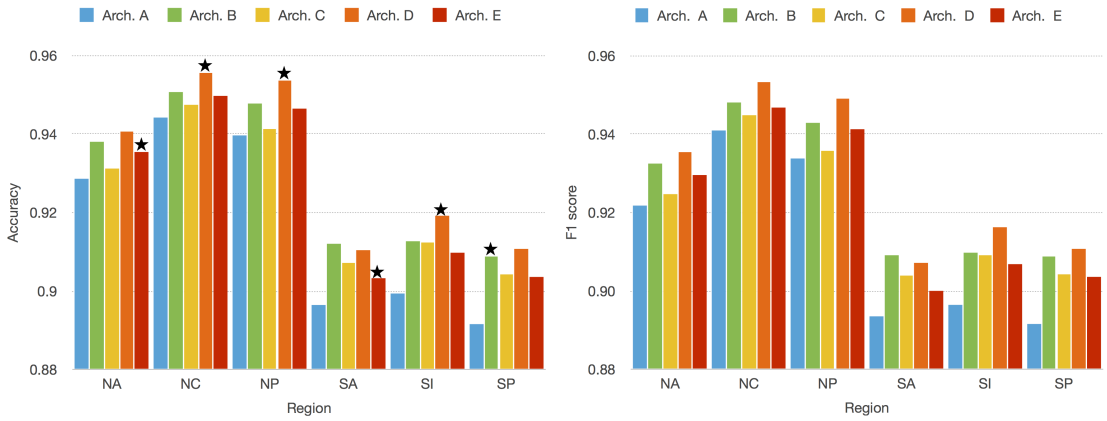
The number of channels, here 40, corresponds to 5 variables at 8 different pressure levels. Location of regions and extent are based on the studies of Wiedenmann et al. [174]. We can distinguish six regions on the globe: the North Pacific Region (NP); the North Atlantic Region (NA); the North Continental Region (NC); the South Pacific Region (SP); the South Atlantic Region (SA), and the South Indian Ocean Region (SI), as shown in Figure 4.5. Each region is roughly centered over a local maximum of ABs main frequency occurrence. We use only images of mid-latitudes regions of both hemispheres on the globe.

The positive class label is assigned to a image if it contains a compact binary blob of size greater or equal to the average AB blob size. The negative class label is assigned to a image if it does not have a binary blob.

After all data pre-processing steps, the generated labelled dataset has approximately 140K images of both ABs and non-ABs, where the number of samples per class is almost balanced. For the regression problem we select the images containing ABs. Then we calculate the centroid or the mass center of a binary blob in the region and a radius of a minimum bounding circular box.

4.2.2 Classification and Localisation Results

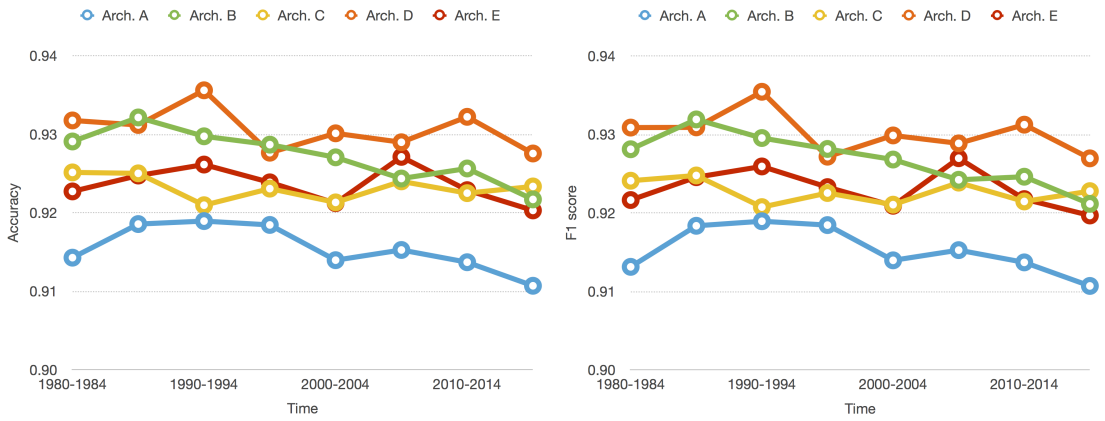
Figure 4.6 (a) and Figure 4.6 (b) show a bar chart of classification accuracy and a bar chart of the mean F1 score of both classes, respectively. The accuracy and F1 values are computed for five architectures described in Table 4.1 and for six regions over the globe, shown in Figure 4.5.



(a) (b)

Figure 4.6: Classification performance of convolutional neural network (CNN) architectures: architecture A, architecture B, architecture C, architecture D, and architecture E; for regions of the North Atlantic Region (NA), the North Continental Region (NC), the North Pacific Region (NP), the South Atlantic Region (SA), the South Indian Ocean Region (SI) and the South Pacific Region (SP). Left bar chart (a) illustrates classification accuracy for each architecture per region and the right chart (b) displays F1 score for each architecture per region [4]. The $\star$ symbol stands for p-value $\ll 0.05$.

In the bar charts, there is the steep decrease in the accuracy and F1 score values of the architectures in the regions of the Southern Hemisphere (i.e., SA, SI, and SP) in comparison to the accuracy and F1 score values of the architectures in the regions of the Northern Hemisphere (i.e., NA, NC, and NP). The bar charts reveal that the architecture D outperforms the other architectures in regions of both hemispheres, regarding the accuracy and F1 score values. In the SA region values of performance metrics are slightly higher for the architecture B than the architecture D. It can be seen that the architecture A achieved the lowest performance for six regions. In general, we can observe a statistically significant increase in the accuracy values (p-value $\ll 0.05$) for: the architecture E in the NA region, the architecture D in the NC region, the architecture D in the NP region, the architecture



(a) (b)

Figure 4.7: Classification performance of Convolutional Neural Network (CNN) architectures: architecture A, architecture B, architecture C, architecture D, and architecture E. Left bar chart (a) illustrates the mean classification accuracy of each architecture over five year periods and the right chart (b) displays the mean F1 score for each architecture over five year periods [4].

E in the SA region, the architecture D in the SI region, and the architecture B in the SP region.

Figure 4.7 shows the mean values of accuracy and F1 score (here, the average of ARs and non-ARs) for the whole globe over time (i.e., averages of five year periods). If we look at the trends over time, we can see the fluctuation in the ACC and F1 score values for all architectures. Although, values of the scores do not level off, we can observe that for some architectures (i.e., architecture B and architecture A) the scores drop systematically over time after the initial steep increase. Moreover, we note that architecture B, architecture C, architecture D, and architecture E converge slightly at similar values of accuracy and F1 score for most of the time.

A bar chart in Figure 4.8 presents concordance correlation coefficients for all models described in Table 4.2 and for regions shown in Figure 4.5. In this chart, there is a steep decline in CCC values for the architectures in regions of the Southern Hemisphere (i.e., SA, SI, and SP) in comparison to CCC values for the architectures in regions of the Northern Hemisphere (i.e., NA, NC, and NP). The chart shows the highest CCC values for the architecture A in regions of both hemispheres, except for the SA region. In contrast, it can be observed that CCC values for the architecture E are the lowest for all six regions.

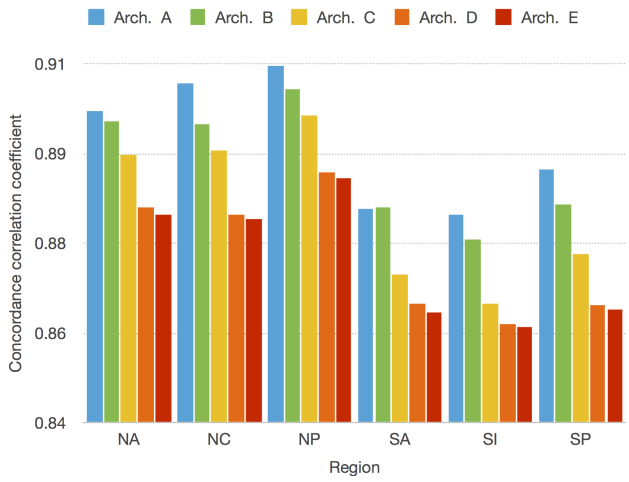


Figure 4.8: Performance of Convolutional Neural Network (CNN) architectures: architecture A, architecture B, architecture C, architecture D, and architecture E; for the North Atlantic Region (NA), the North Continental Region (NC), the North Pacific Region (NP), the South Atlantic Region (SA), the South Indian Ocean Region (SI) and the South Pacific Region (SP). The bar chart illustrates the concordance correlation coefficient for each architecture per region [4].

Overall, we can see that the more complex an architecture becomes, the more the CCC values decrease.

Figure 4.9 (a), (b), and (c) display bar charts of mean percentage error of the three predicted parameters describing the location of AB phenomena for all architectures described in Table 4.2 and for the six regions shown in Figure 4.5. Figure 4.9 (a) shows the MPE of the latitudinal mass centre position of ABs. It can be seen that all CNN architectures overestimate the latitudinal parameter (positive error) in all six regions. The bar chart also reveals that there has been a steep decrease in overestimating the parameter by all architectures in the NP region and the SI region. What can be observed in the chart is that the architecture E has the highest MPE in four out of six regions (i.e., NA, SA, SI, and SP). The MPE of the latitudinal parameter decreases statistically significantly (p-value $\ll 0.05$) for: the architecture A in the SA region, the architecture B in the SP region, the architecture C in the NA region, and the architecture D in the NC, NP and SI regions.

Figure 4.9 (b) shows the MPE of the longitudinal position of AB phenomena. It can be seen that different architectures underestimate the longitudinal parameter in five out

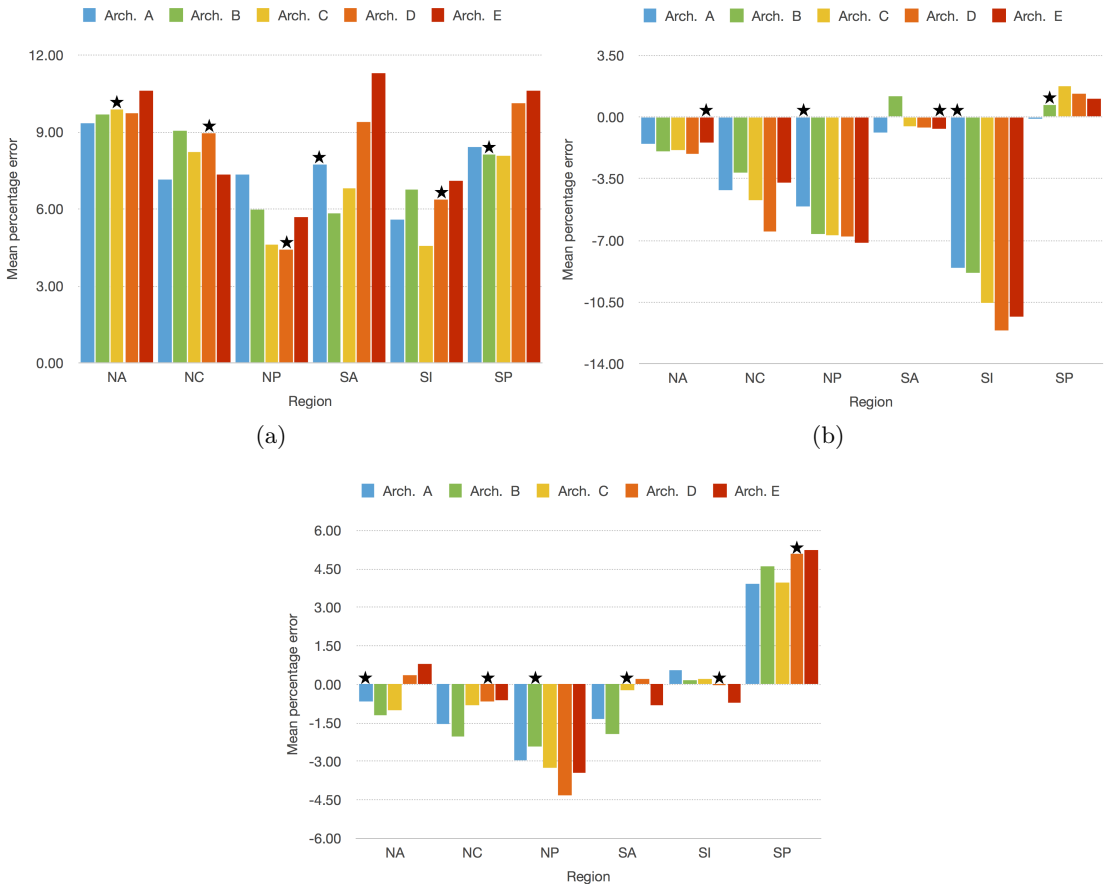


Figure 4.9: Performance of Convolutional Neural Network (CNN) architectures: architecture A, architecture B, architecture C, architecture D, and architecture E; for regions of the North Atlantic Region (NA), the North Continental Region (NC), the North Pacific Region (NP), the South Atlantic Region (SA), the South Indian Ocean Region (SI) and the South Pacific Region (SP). Top left bar chart (a) illustrates mean percentage error for each architecture in estimating the latitudinal position of the mass centre of atmospheric blocks (ABs) per region and the top right chart (b) displays mean percentage error for each architecture in estimating the longitudinal position of the mass centre of ABs per region. The bottom chart shows mean percentage error for each architecture in estimating the radius of the mass centre of ABs per region [4]. The $\star$ symbol stands for a p-value $\ll 0.05$.

of six regions (i.e., NA, NC, NP, SA, and SI). The exception to this is the architecture B in the region SA that the MPE value is overestimated. The chart also indicates that MPE values are positive for most architectures in one out of six regions (i.e., the SP region). Architecture A is the only exception, which has a small negative MPE value. The MPE of longitudinal parameter undergoes a statistically significant decrease (p-value $\ll 0.05$) for: the architecture A in the NP and SI regions; the architecture B in the SP region, and the architecture E in the NA and SA regions. The NC region is the exception to this in which there is not a statistically significant decrease in the MPE values.

Figure 4.9 (c) displays the MPE of a radius of a minimum bounding box located in the centre of ABs. We can observe low MPE values for the architectures in four out of six regions (i.e., NA, NC, SA, and SI). Furthermore, it can be observed that the MPE values for all architectures are high in NP and SP regions. Overall, the bar chart reveals that the architectures underestimate the radius in five out of six regions. The MPE of the radius undergoes a statistically significant decreases (p-value $\ll 0.05$) for: the architecture A in the NA region, the architecture B in the NP region, the architecture C in the SA region, and the architecture D in the NC and SP regions.

4.3 Discussion and Summary

In this chapter, we proposed a hierarchical pattern recognition method for identifying atmospheric blocking phenomena in climate reanalysis products. The method consists of two stages: the detection stage that is formulated as a binary classification and the localisation stage that is defined as a regression task. We trained five different architectures of a generic CNN-based classifier and CNN-based regressor respectively designed for each stage of the AB pattern recognition method. The ground truth labels were generated by a heuristic identification method of ABs. We also notice that CNNs can be biased by the ground truth data produced by the heuristic identification method using the threshold-based criteria to detect ABs. We note that characterising the influence of using different ground truth data is beyond the scope of this thesis.

We demonstrated that the proposed pattern recognition method accurately classifies AB phenomena in the ERA-Interim reanalysis data by using CNNs. The classification accuracy and F1 score values are the highest for all architectures in all three regions of the Northern Hemisphere (NH). That can be justified by the fact that AB phenomena are much more frequent in the NH than in the Southern Hemisphere (SH). Moreover, ABs have

important consequences on the weather of densely populated Europe and North America, that is why innumerable observational data improve the ability of reanalysis to represent ABs in the NH correctly. Furthermore, the values of the mean accuracy and F1 score fluctuate over time. This variability in the values may be explained by the number of occurrence of ABs, which can be sometimes modulated by other periodic climate regimes, for example, the North Atlantic Oscillation [175, 176]. We also observe that the values of accuracy and F1 score statistically significantly increase for CNN architectures with a large number of parameters for various regions on the globe.

Another observation is that the values of the concordance correlation coefficient are generally high for all CNN architectures in all regions of the NH. However, that is not the case for the architectures in the SH. That may also be partially explained by better observational characteristics of ABs in the NH and more improvements in the ability of the reanalysis to represent atmospheric dynamics of these phenomena correctly.

The values of the MPE of the latitudinal mass centre position of ABs are usually overestimated. On ther other hand, the MPE values of the longitudinal mass center position of these phenomena are usually underestimated. There are two reasons for this. ABs occur in certain longitudinal and latitudinal localisation in the NH, and these phenomena often occur at slightly lower latitude positions in the SH than in the NH. It can be observed that the values of the MPE statistically significantly decrease for CNNs without any specific trend in particular geographical areas. Moreover, the values of the MPE of a radius of a minimum bounding box located in the mass centre of ABs are low in most of the regions. That can suggest that despite a wide spectrum of spatial location, ABs are probably more similar in size in both hemispheres. Moreover, the MPE values are much higher in the Pacific Ocean regions, which could suggest that the reanalysis may have problems with representing ABs spatial size due to the vast extent of the ocean.

Chapter 5

Conclusion and Future Work

This chapter concludes this thesis. We first provide a summary of the research outcomes discussed in each of the previous chapters. Then, we respond to the research questions formulated in Section 1.5.1. We subsequently discuss the limitations of our work and what has been learnt while carrying out this research. Finally, we propose future research directions in the area of weather pattern recognition.

5.1 Research Outcomes

In this thesis, we focused on the identification of weather phenomena in climate data. We aimed at developing and applying pattern recognition methods for identifying the Atmospheric River (AR) and Atmospheric Blocking (AB) phenomena.

In Chapter 3, we developed and explored an automated method for recognizing ARs that combines topological data analysis with machine learning methods. The key advantage of the topological feature descriptors used in our method is that it is a threshold-free method (i.e., no threshold requirements on physical variables) that succinctly encapsulates the topological structure of ARs. In other words, there is no threshold parameter retuning, unlike engineered heuristic approaches used by other objective AR identification methods. Moreover, we demonstrated that the proposed method is reliable, robust and performs well by testing it on a wide range of spatial and temporal resolutions of the Community Atmosphere Model v5.1 (CAM5.1) outputs. Furthermore, despite background noise, low-intensity AR signals and the existence of other events within the 2-D snapshots, our method is shown to work accurately. The method better performed for low-resolution data. We

speculated that high-resolution simulations produce noisy spatial patterns that mislead machine learning models more easily than low-resolution simulations. Also, we made a comparison of our method to other AR identification methods when applied to reanalysis data. We observed that our method tended to misclassify some ARs. In the case of the 2nd Modern-Era Retrospective Analysis for Research and Applications (MERRA-2) reanalysis data, the typical failure modes are not different than for the CAM5.1 datasets. Our method frequently produces false negatives in recognizing ARs in the MERRA-2 data. It is known that the reanalysis data may have higher background noise and low-intensity AR signals. That would result in higher frequency in the more active storm tracks, which may explain the rapid drop-off in AR detection associated with this method. It is also worth mentioning that our method excels at lower latitudes, where it identifies a relatively larger number of ARs than most AR identification methods. As we noted before, our method typically identifies fewer ARs than other methods used in the intercomparison. What is more, the ARs identified by our method have shorter average durations than for other AR identification methods.

In Chapter 4, we explored a pattern recognition method for classifying and localising AB phenomena based on deep Convolutional Neural Networks (CNNs). We demonstrated that the proposed method accurately detects AB phenomena in the European Centre for Medium-Range Weather Forecasts Interim Reanalysis (ERA-Interim) data. The proposed CNN architectures tend to achieve better AB classification and localisation performances in regions of the Northern Hemisphere (NH) than in regions of the Southern Hemisphere (SH). That can be justified by the fact that AB phenomena are much more frequent in the NH than in the SH. Moreover, ABs have important consequences on the weather of densely populated areas of Europe and North America. That is why innumerable observational data improve the ability of reanalysis product to represent ABs in the NH correctly. Furthermore, we observed that the classification accuracy values fluctuated over time. This variability in the values may be explained by the number of AB occurrences that is sometimes modulated by other periodic climate regimes, such as the North Atlantic Oscillation [175, 176]. We also indicate that CNNs performance significantly increases for the architectures with a large number of parameters for various regions on the globe. CNNs achieved the lower estimation error of ABs localisation in regions of the NH than in regions of the SH. That can be caused by two reasons. ABs tend to occur in certain longitudinal and latitudinal localisation in the NH, and these phenomena often occur at slightly lower latitude positions in the SH than in the NH. Moreover, the estimation error values are much higher

in the Pacific Ocean region than in other regions. That could suggest that the reanalysis product may have problems with correctly representing ABs spatial size because of the vast extent of the Pacific ocean.

5.2 Responding to Research Questions

We formulated three main research questions. Each question consists of three sub-questions described in Section 1.5.1. Below we respond to each of these questions in turn based on the findings presented in Chapter 3 and 4:

- The dual-step method for AR pattern recognition based on topological data analysis and machine learning can automatically discover regularities in raw climate data to identify AR phenomena. This is the first such a method that uses topological feature descriptors to characterise ARs. The method is complementary to other AR identification methods. (RQ1: RQ1.a; RQ1.b, RQ2: RQ2.a)

- The experimental results indicate that the AR pattern recognition method can provide a new way of computing the climatological statistics of AR phenomena. (RQ1: RQ1.b; RQ1.c)

- Topological feature descriptors used in our threshold-free feature extraction succinctly encapsulate the most important structural characteristics of ARs. Machine learning model benefits from these topological descriptors that is demonstrated by the high classification accuracy. (RQ2: RQ2.a; RQ2.b)

- The experimental results show that the performance of AR recognition methods is sensitive to changes in spatiotemporal resolution of climate model simulation outputs, i.e. the CAM5.1. (RQ2: RQ2.c)

- The hierarchical AB pattern recognition method based on state-of-the-art deep CNN architectures can automatically extract and detect ABs in climate reanalysis data in different regions over the globe. (RQ1: RQ1.a; RQ1.c)

- The experimental results demonstrate that the proposed CNN architectures achieve better AB detection and localization performances in regions of the Northern Hemisphere than in regions of the Southern Hemisphere. (RQ1: RQ1.b, RQ3: RQ3.a; RQ3.b;)

- The AB detection performance significantly increases for the deep CNN architectures. In contrast, the estimation error of AB location significantly decreases in the localisation task for the shallow CNN architectures. (RQ3: RQ3.c)

5.3 Lessons Learnt and Limitations

Here, we summarise lessons learnt and limitations while carrying out research on weather phenomenon recognition:

- Modern pattern recognition methods can be complementary to heuristic identification methods of weather phenomena. However, the abstract representation of weather phenomena is learned by machine learning models, only limited understanding of the mechanisms responsible for changes in weather phenomena can be provided.

- Classification results can be biased by the ground truth data produced by heuristic (hand-crafted) identification methods using the threshold-based criteria to detect weather phenomena, such as ARs and ABs.

- Collecting ground truth is one of the main challenges in weather phenomenon recognition tasks. That is why designing protocols and platforms for hand-labelling campaigns can be beneficial.

- Low-intensity AR signals and the existence of other events within the 2-D snapshots can cause problems for topological descriptors to correctly summarise AR characteristics.

- Our AR recognition method tends to better perform for low-resolution data. We speculate that this is because high-resolution simulations produce noisy spatial patterns that mislead machine learning models more easily than low-resolution simulations.

- Deep CNN architectures are able to achieve high classification performance despite a low number of instances. As the evidence, we observe that the values of evaluation metrics, such as classification accuracy, significantly increase for deep CNNs with a large number of parameters for various regions on the globe.

- The high estimation error of AB location measured by the mean percentage error in the Pacific Ocean region suggests that the climate reanalysis data do not correctly

represent ABs spatial statistics in the region. We hypothesise that this is caused by climate reanalysis disability to represent atmospheric conditions at sufficient spatial resolution.

5.4 Future Work

All the outcomes of studies presented in this thesis suggest that pattern recognition for weather phenomena is a complex and challenging task. Our studies demonstrated that providing knowledge of atmospheric phenomena, such as ARs or ABs, can be beneficial for designing pattern recognition methods. Moreover, a vast amount of climate data are rapidly growing with advances in high performance computing technology. That will require efficient and reliable post-processing in the future. For those reasons, there are several research directions that these studies can be pursued in the future.

Characterising the influence of using different ground truth data on pattern recognition methods performance could help understand biases in the training data and design weather phenomenon recognition methods. That will allow the climate science community to understand the real capabilities of pattern recognition methods that are based on machine learning and deep learning methods. However, that would be of interest to the climate science community; there is a lack of labelled climate data to train these methods. To overcome that limitation transfer learning between different weather phenomenon recognition tasks could be investigated. For example, pre-trained machine learning methods on one phenomenon recognition task can be applied to another recognition task.

Furthermore, the creation of annotation collection platforms is needed because there is a large variance of weather phenomenon indices among hand-crafted weather identification methods. Some weather phenomena are complicated to detect since they are very complex; for instance, weather fronts [79]. The existence of the one absolute ground truth is questionable, and it should be replaced by data collected from labelling campaigns, for example, ClimateNet [177].

Testing weather phenomenon recognition methods on direct observations via satellite images and coupled climate model simulations (e.g., CMIP5 [35] or CMIP6 [83]) could help the climate community understand the difference in climatologies of various climate models. Moreover, weather phenomenon recognition methods could be an objective metric for evaluating both reanalysis products and climate models against observational data.

Exploring more advanced deep learning methods could facilitate analysis of ever-increasing

climate datasets. For instance, developing 3D CNN architectures to identify weather phenomena would help analyse more complex spatial climate data dependencies. Also, the application of convolutional recurrent neural networks (e.g., CNN combined with LSTM [78]) that explore spatiotemporal dependencies in climate data would possibly reveal more about dynamical and statistical features of weather phenomena, such as AR and AB phenomena.

Current machine learning approaches to weather pattern recognition provide novel and powerful ways of accurately and efficiently identifying weather phenomena in climate data. Nevertheless, machine learning neither obeys the fundamental laws of physical systems nor incorporates domain knowledge. As a consequence, weather pattern recognition methods do not generalise well to unseen scenarios. That is why the challenge of incorporating atmospheric physics and domain knowledge into machine learning models still remains.

Bibliography

[1] G. Muszynski, K. Kashinath, V. Kurlin, M. Wehner, and Prabhat. Topological data analysis and machine learning for recognizing atmospheric river patterns in large climate datasets. *Geoscientific Model Development*, 12(2):613–628, 2019.

[2] C. A. Shields, J. J. Rutz, L.-Y. Leung, F. M. Ralph, M. Wehner, B. Kawzenuk, J. M. Lora, E. McClenny, T. Osborne, A. E. Payne, P. Ullrich, A. Gershunov, N. Goldenson, B. Guan, Y. Qian, A. M. Ramos, C. Sarangi, S. Sellars, I. Gorodetskaya, K. Kashinath, V. Kurlin, K. Mahoney, G. Muszynski, R. Pierce, A. C. Subramanian, R. Tome, D. Waliser, D. Walton, G. Wick, A. Wilson, D. Lavers, Prabhat, A. Collow, H. Krishnan, G. Magnusdottir, and P. Nguyen. Atmospheric river tracking method intercomparison project (artmip): project goals and experimental design. *Geoscientific Model Development*, 11(6):2455–2474, 2018.

[3] Jonathan J Rutz, Christine A Shields, Juan M Lora, Ashley E Payne, Bin Guan, Paul Ullrich, Travis O'Brien, L Ruby Leung, F Martin Ralph, Michael Wehner, et al. The atmospheric river tracking method intercomparison project (artmip): quantifying uncertainties in atmospheric river climatology. *Journal of Geophysical Research: Atmospheres*, 2019.

[4] Grzegorz Muszynski, Prabhat Mr, Jan Balewski, Karthik Kashinath, Wehner Michael, and Vitaliy Kurlin. Atmospheric blocking pattern recognition in global climate model simulation data. In *25th International Conference on Pattern Recognition (ICPR)*. IEEE, 2021.

[5] IPCC Climate Change et al. The physical science basis. 2007.

[6] Gabriele C Hegerl, Stefan Brönnimann, Andrew Schurer, and Tim Cowan. The early

20th century warming: anomalies, causes, and consequences. *Wiley Interdisciplinary Reviews: Climate Change*, 9(4):e522, 2018.

[7] Dahe Qin, GK Plattner, M Tignor, SK Allen, J Boschung, A Nauels, Y Xia, V Bex, PM Midgley, et al. Climate change 2013: the physical science basis. *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (eds TF Stocker et al.)*, pages 5–14, 2014.

[8] Svante Arrhenius. Xxxi. on the influence of carbonic acid in the air upon the temperature of the ground. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(251):237–276, 1896.

[9] Clara Deser, Michael A Alexander, Shang-Ping Xie, and Adam S Phillips. Sea surface temperature variability: Patterns and mechanisms. *Annual review of marine science*, 2:115–143, 2010.

[10] Syukuro Manabe and Ronald J Stouffer. Sensitivity of a global climate model to an increase of co2 concentration in the atmosphere. *Journal of Geophysical Research: Oceans*, 85(C10):5529–5554, 1980.

[11] Marika M Holland and Cecilia M Bitz. Polar amplification of climate change in coupled models. *Climate Dynamics*, 21(3-4):221–232, 2003.

[12] Doug M Smith, James A Screen, Clara Deser, Judah Cohen, John C Fyfe, Javier García-Serrano, Thomas Jung, Vladimir Kattsov, Daniela Matei, Rym Msadek, et al. The polar amplification model intercomparison project (pamip) contribution to cmip6: investigating the causes and consequences of polar amplification. 2019.

[13] Mark C Serreze and Roger G Barry. Processes and impacts of arctic amplification: A research synthesis. *Global and planetary change*, 77(1-2):85–96, 2011.

[14] Jennifer A Francis and Stephen J Vavrus. Evidence linking arctic amplification to extreme weather in mid-latitudes. *Geophysical research letters*, 39(6), 2012.

[15] James A Screen and Ian Simmonds. Exploring links between arctic amplification and mid-latitude weather. *Geophysical Research Letters*, 40(5):959–964, 2013.

[16] John E Walsh. Intensified warming of the arctic: Causes and impacts on middle latitudes. *Global and Planetary Change*, 117:52–63, 2014.

[17] Matthew Collins, Reto Knutti, Julie Arblaster, Jean-Louis Dufresne, Thierry Fichefet, Pierre Friedlingstein, Xuejie Gao, William J Gutowski, Tim Johns, Gerhard Krinner, et al. Long-term climate change: projections, commitments and irreversibility. In *Climate Change 2013-The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 1029–1136. Cambridge University Press, 2013.

[18] David G Vaughan, Gareth J Marshall, William M Connolley, Claire Parkinson, Robert Mulvaney, Dominic A Hodgson, John C King, Carol J Pudsey, and John Turner. Recent rapid regional climate warming on the antarctic peninsula. *Climatic change*, 60(3):243–274, 2003.

[19] Alexandra Jahn, Jennifer E Kay, Marika M Holland, and David M Hall. How predictable is the timing of a summer ice-free arctic? *Geophysical Research Letters*, 43(17):9113–9120, 2016.

[20] MC Serreze, AP Barrett, JC Stroeve, DN Kindig, and MM Holland. The emergence of surface-based arctic amplification. *The Cryosphere*, 3(1):11, 2009.

[21] Isaac M Held and Brian J Soden. Robust responses of the hydrological cycle to global warming. *Journal of climate*, 19(21):5686–5699, 2006.

[22] Gerard A Meehl, Thomas F Stocker, William D Collins, Pierre Friedlingstein, T Gaye, Jonathan M Gregory, Akio Kitoh, Reto Knutti, James M Murphy, Akira Noda, et al. Global climate projections. 2007.

[23] James E Overland and Muyin Wang. Large-scale atmospheric circulation changes are associated with the recent loss of arctic sea ice. *Tellus A: Dynamic Meteorology and Oceanography*, 62(1):1–9, 2010.

[24] Jennifer A Francis and Stephen J Vavrus. Evidence for a wavier jet stream in response to rapid arctic warming. *Environmental Research Letters*, 10(1):014005, 2015.

[25] Ralf Jaiser, Klaus Dethloff, and Dörthe Handorf. Stratospheric response to arctic sea ice retreat and associated planetary wave propagation changes. *Tellus A: Dynamic Meteorology and Oceanography*, 65(1):19375, 2013.

[26] Ebrahim Nabizadeh, Pedram Hassanzadeh, Da Yang, and Elizabeth A Barnes. Size of the atmospheric blocking events: Scaling law and response to climate change. *Geophysical Research Letters*, 46(22):13488–13499, 2019.

[27] Tim Woollings, Ben Harvey, and Giacomo Masato. Arctic warming, atmospheric blocking and cold european winters in cmip5 models. *Environmental Research Letters*, 9(1):014002, 2014.

[28] Hyo-Seok Park, Sukyoung Lee, Seok-Woo Son, Steven B Feldstein, and Yu Kosaka. The impact of poleward moisture and sensible heat flux on arctic winter sea ice variability. *Journal of Climate*, 28(13):5030–5040, 2015.

[29] Matthew Newman, George N Kiladis, Klaus M Weickmann, F Martin Ralph, and Prashant D Sardeshmukh. Relative contributions of synoptic and low-frequency eddies to time-mean atmospheric moisture transport, including the role of atmospheric rivers. *Journal of climate*, 25(21):7341–7361, 2012.

[30] Chengji Liu and Elizabeth A Barnes. Extreme moisture transport into the arctic linked to rossby wave breaking. *Journal of Geophysical Research: Atmospheres*, 120(9):3774–3788, 2015.

[31] Wilhelm Kirch, Bettina Menne, and Roberto Bertollini. Extreme weather events and public health responses. 2005.

[32] US De, Rajiv Kumar Dube, and GS Prakasa Rao. Extreme weather events over india in the last 100 years. *J. Ind. Geophys. Union*, 9(3):173–187, 2005.

[33] Bogdan Antonescu, David M Schultz, Hugo MAM Ricketts, and Dragoş Ene. Theories on tornado and waterspout formation in ancient greece and rome. *Weather, Climate, and Society*, 11(4):889–900, 2019.

[34] Peter Bauer, Alan Thorpe, and Gilbert Brunet. The quiet revolution of numerical weather prediction. *Nature*, 525(7567):47–55, 2015.

[35] Karl E Taylor, Ronald J Stouffer, and Gerald A Meehl. An overview of cmip5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498, 2012.

[36] Michael F Wehner, Kevin A Reed, Fuyu Li, Julio Bacmeister, Cheng-Ta Chen, Christopher Paciorek, Peter J Gleckler, Kenneth R Sperber, William D Collins, Andrew Gettelman, et al. The effect of horizontal resolution on simulation quality in the community atmospheric model, cam5. 1. *Journal of Advances in Modeling Earth Systems*, 6(4):980–997, 2014.

[37] BJ Hoskins. Dynamical processes in the atmosphere and the use of models. *Quarterly Journal of the Royal Meteorological Society*, 109(459):1–21, 1983.

[38] American Meteorological Society. Weather System. Glossary of Meteorology. [Available online at $http://glossary.ametsoc.org/wiki/Weather\_system$]. 2015.

[39] Reginald E Newell, Nicholas E Newell, Yong Zhu, and Courtney Scott. Tropospheric rivers?–a pilot study. *Geophysical Research Letters*, 19(24):2401–2404, 1992.

[40] Reginald E Newell and Yong Zhu. Tropospheric rivers: A one-year record and a possible application to ice core data. *Geophysical research letters*, 21(2):113–116, 1994.

[41] Yong Zhu and Reginald E Newell. Atmospheric rivers and bombs. *Geophysical Research Letters*, 21(18):1999–2002, 1994.

[42] Yong Zhu and Reginald E Newell. A proposed algorithm for moisture fluxes from atmospheric rivers. *Monthly weather review*, 126(3):725–735, 1998.

[43] F Martin Ralph, Paul J Neiman, and Gary A Wick. Satellite and caljet aircraft observations of atmospheric rivers over the eastern north pacific ocean during the winter of 1997/98. *Monthly Weather Review*, 132(7):1721–1745, 2004.

[44] AMS. Atmospheric River. Glossary of Meteorology. [Available online at $http://glossary.ametsoc.org/wiki/Atmospheric\_river$]. 2019.

[45] F Martin Ralph, Michael D Dettinger, Mary M Cairns, Thomas J Galarneau, and John Eylander. Defining "atmospheric river": How the glossary of meteorology helped resolve a debate. *Bulletin of the American Meteorological Society*, 99(4):837–839, 2018.

[46] David A Lavers and Gabriele Villarini. The nexus between atmospheric rivers and extreme precipitation across europe. *Geophysical Research Letters*, 40(12):3259–3264, 2013.

[47] Bin Guan, Noah P Molotch, Duane E Waliser, Eric J Fetzer, and Paul J Neiman. Extreme snowfall events linked to atmospheric rivers and surface air temperature via satellite measurements. *Geophysical Research Letters*, 37(20), 2010.

[48] FM Ralph, T Coleman, PJ Neiman, RJ Zamora, and MD Dettinger. Observed impacts of duration and seasonality of atmospheric-river landfalls on soil moisture and runoff in coastal northern california. *Journal of Hydrometeorology*, 14(2):443–459, 2013.

[49] David A Lavers, Florian Pappenberger, and Ervin Zsoter. Extending medium-range predictability of extreme hydrological events in europe. *Nature communications*, 5(1):1–7, 2014.

[50] Maximiliano Viale and Mario N Nuñez. Climatology of winter orographic precipitation over the subtropical central andes and associated synoptic and regional characteristics. *Journal of Hydrometeorology*, 12(4):481–507, 2011.

[51] RC Blamey, AM Ramos, RM Trigo, R Tomé, and CJC Reason. The influence of atmospheric rivers over the south atlantic on winter rainfall in south africa. *Journal of Hydrometeorology*, 19(1):127–142, 2018.

[52] Alexandre M Ramos, Ricardo Tomé, Ricardo M Trigo, Margarida LR Liberato, and Joaquim G Pinto. Projected changes in atmospheric rivers affecting europe in cmip5 models. *Geophysical Research Letters*, 43(17):9315–9323, 2016.

[53] Bin Guan and Duane E Waliser. Atmospheric rivers in 20 year weather and climate simulations: A multimodel, global evaluation. *Journal of Geophysical Research: Atmospheres*, 122(11):5556–5581, 2017.

[54] Daniel F Rex. Blocking action in the middle troposphere and its effect upon regional climate: I. an aerological study of blocking action. *Tellus*, 2(3):196–211, 1950.

[55] DANIEL F REX. Blocking action in the middle troposphere and its effect upon regional climate: Ii. the climatology of blocking action. *Tellus*, 2(4):275–301, 1950.

[56] C-G Rossby. On the dynamics of certain types of blocking waves. *Journal of the Chinese Geophysical Society*, 2:1–13, 1951.

[57] Ricardo García-Herrera, José Díaz, Ricardo M Trigo, Jürg Luterbacher, and Erich M Fischer. A review of the european summer heat wave of 2003. *Critical Reviews in Environmental Science and Technology*, 40(4):267–306, 2010.

[58] Tania Buehler, Christoph C Raible, and Thomas F Stocker. The relationship of winter season north atlantic blocking frequencies to extreme cold or dry spells in the era-40. *Tellus A: Dynamic Meteorology and Oceanography*, 63(2):174–187, 2011.

[59] Daniel L Swain, Michael Tsiang, Matz Haugen, Deepti Singh, Allison Charland, Bala Rajaratnam, and Noah S Diffenbaugh. The extraordinary california drought of 2013/2014: Character, context, and the role of climate change. *Bull. Am. Meteorol. Soc*, 95(9):S3–S7, 2014.

[60] Chi-Cherng Hong, Huang-Hsiung Hsu, Nai-Hsin Lin, and Hsun Chiu. Roles of european blocking and tropical-extratropical interaction in the 2010 pakistan flooding. *Geophysical Research Letters*, 38(13), 2011.

[61] Li Dong, Chandana Mitra, Seth Greer, and Ethan Burt. The dynamical linkage of atmospheric blocking to drought, heatwave and urban heat island in southeastern us: A multi-scale case study. *Atmosphere*, 9(1):33, 2018.

[62] Stefano Tibaldi and Franco Molteni. Atmospheric blocking in observation and models. In *Oxford Research Encyclopedia of Climate Science*. 2018.

[63] AMS. Blocking. Glossary of Meteorology. [Available online at *https* : *//glossary.ametsoc.org/wiki/Blocking*]. 2019.

[64] S Pfahl and Heini Wernli. Quantifying the relevance of atmospheric blocking for co-located temperature extremes in the northern hemisphere on (sub-) daily time scales. *Geophysical Research Letters*, 39(12), 2012.

[65] Michael D Dettinger, Fred Martin Ralph, Tapash Das, Paul J Neiman, and Daniel R Cayan. Atmospheric rivers, floods and the water resources of california. *Water*, 3(2):445–478, 2011.

[66] Jana Sillmann, Mischa Croci-Maspoli, Malaak Kallache, and Richard W Katz. Extreme cold winter temperatures in europe under the influence of north atlantic atmospheric blocking. *Journal of Climate*, 24(22):5899–5913, 2011.

[67] Gotzon Gangoiti, Lucio Alonso, Marino Navazo, Amaia Albizuri, Gorka Perez-Landa, Monica Matabuena, Veronica Valdenebro, Mercedes Maruri, José Antonio García, and Millán M Millán. Regional transport of pollutants over the bay of biscay: analysis of an ozone episode under a blocking anticyclone in west-central europe. *Atmospheric Environment*, 36(8):1349–1361, 2002.

[68] Intergovernmental Panel On Climate Change. *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2014.

[69] Kevin I Hodges. A general method for tracking analysis and its application to meteorological data. *Monthly Weather Review*, 122(11):2573–2586, 1994.

[70] KI Hodges. Feature tracking on the unit sphere. *Monthly Weather Review*, 123(12):3458–3465, 1995.

[71] Hans Alexandersson. A homogeneity test applied to precipitation data. *Journal of climatology*, 6(6):661–675, 1986.

[72] Thomas C Peterson, David R Easterling, Thomas R Karl, Pavel Groisman, Neville Nicholls, Neil Plummer, Simon Torok, Ingeborg Auer, Reinhard Boehm, Donald Gullett, et al. Homogeneity adjustments of in situ atmospheric climate data: a review. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 18(13):1493–1517, 1998.

[73] Jonathan T Overpeck, Gerald A Meehl, Sandrine Bony, and David R Easterling. Climate data challenges in the 21st century. *science*, 331(6018):700–702, 2011.

[74] James D Ford, Simon E Tilleard, Lea Berrang-Ford, Malcolm Araos, Robbert Biesbroek, Alexandra C Lesnikowski, Graham K MacDonald, Angel Hsu, Chen Chen, and Livia Bizikova. Opinion: Big data has big potential for applications to climate change adaptation. *Proceedings of the National Academy of Sciences*, 113(39):10729–10732, 2016.

[75] Dick P Dee, S M Uppala, AJ Simmons, Paul Berrisford, P Poli, S Kobayashi, U Andrae, MA Balmaseda, G Balsamo, d P Bauer, et al. The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, 137(656):553–597, 2011.

[76] Ronald Gelaro, Will McCarty, Max J Suárez, Ricardo Todling, Andrea Molod, Lawrence Takacs, Cynthia A Randles, Anton Darmenov, Michael G Bosilovich, Rolf Reichle, et al. The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of Climate*, 30(14):5419–5454, 2017.

[77] Urs Neu, Mirseid G Akperov, Nina Bellenbaum, Rasmus Benestad, Richard Blender, Rodrigo Caballero, Angela Cocozza, Helen F Dacre, Yang Feng, Klaus Fraedrich, et al. Imilast: A community effort to intercompare extratropical cyclone detection and tracking algorithms. *Bulletin of the American Meteorological Society*, 94(4):529–547, 2013.

[78] Sookyung Kim, Hyojin Kim, Joonseok Lee, Sangwoong Yoon, Samira Ebrahimi Kahou, Karthik Kashinath, and Mr Prabhat. Deep-hurricane-tracker: Tracking and forecasting extreme climate events. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1761–1769. IEEE, 2019.

[79] James C Biard and Kenneth E Kunkel. Automated detection of weather fronts using a deep learning neural network. *Advances in Statistical Climatology, Meteorology and Oceanography*, 5(2):147–160, 2019.

[80] Marielle Cristine Pinheiro, PA Ullrich, and Richard Grotjahn. Atmospheric blocking and intercomparison of objective detection methods: flow field characteristics. *Climate Dynamics*, pages 1–28, 2019.

[81] Christine A Shields and Jeffrey T Kiehl. Atmospheric river landfall-latitude changes in future climate simulations. *Geophysical Research Letters*, 43(16):8775–8782, 2016.

[82] Michael Sprenger, Georgios Fragkoulidis, Hanin Binder, Mischa Croci-Maspoli, Pascal Graf, Christian M Grams, Peter Knippertz, Erica Madonna, Sebastian Schemm, Bojan Škerlak, et al. Global climatologies of eulerian and lagrangian flow features based on era-interim. *Bulletin of the American Meteorological Society*, 98(8):1739–1748, 2017.

[83] Veronika Eyring, Sandrine Bony, Gerald A Meehl, Catherine A Senior, Bjorn Stevens, Ronald J Stouffer, and Karl E Taylor. Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958, 2016.

[84] Grzegorz Muszynski, Vitaliy Kurlin, Dmitriy. Morozov, Michael Wehner, Karthik. Kashinath, and Prabhat. *Big Earth Data Analytics in Earth, Atmospheric and Ocean Sciences*, chapter Topological Methods for Pattern Detection in Climate Data. John Wiley & Sons, 2020.

[85] C. A. Shields, J. J. Rutz, L.-Y. Leung, F. M. Ralph, M. Wehner, B. Kawzenuk, J. M. Lora, E. McClenny, T. Osborne, A. E. Payne, P. Ullrich, A. Gershunov, N. Goldenson, B. Guan, Y. Qian, A. M. Ramos, C. Sarangi, S. Sellars, I. Gorodetskaya, K. Kashinath, V. Kurlin, K. Mahoney, G. Muszynski, R. Pierce, A. C. Subramanian, R. Tome, D. Waliser, D. Walton, G. Wick, A. Wilson, D. Lavers, Prabhat, A. Collow, H. Krishnan, G. Magnusdottir, and P. Nguyen. Atmospheric river tracking method intercomparison project (artmip): project goals and experimental design. *Geoscientific Model Development*, 11(6):2455–2474, 2018.

[86] Tim D Hewson and Helen A Titley. Objective identification, typing and tracking of the complete life-cycles of cyclonic features at high spatial resolution. *Meteorological Applications*, 17(3):355–381, 2010.

[87] Steven J Lambert. A cyclone climatology of the canadian climate centre general circulation model. *Journal of Climate*, 1(1):109–115, 1988.

[88] John W Nielsen and Randall M Dole. A survey of extratropical cyclone characteristics during gale. *Monthly Weather Review*, 120(7):1156–1168, 1992.

[89] Sabine Eckhardt, Andreas Stohl, Heini Wernli, Paul James, Caroline Forster, and Nicole Spichtinger. A 15-year climatology of warm conveyor belts. *Journal of climate*, 17(1):218–237, 2004.

[90] Tim D Hewson. Objective fronts. *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, 5(1):37–65, 1998.

[91] J Jenkner, M Sprenger, I Schwenk, C Schwierz, S Dierer, and D Leuenberger. Detection and climatology of fronts in a high-resolution model reanalysis over the alps.

Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling, 17(1):1–18, 2010.

[92] Patrick Koch, Heini Wernli, and Huw C Davies. An event-based jet-stream climatology and typology. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 26(3):283–301, 2006.

[93] Sebastian Limbach, Elmar Schömer, and Heini Wernli. Detection, tracking and event localization of jet stream features in 4-d atmospheric data. *Geoscientific Model Development*, 5(2):457–470, 2012.

[94] Jennifer Louise Catto and S Pfahl. The importance of fronts for extreme precipitation. *Journal of Geophysical Research: Atmospheres*, 118(19):10–791, 2013.

[95] Sebastian Schemm and Michael Sprenger. Frontal-wave cyclogenesis in the north atlantic–a climatological characterisation. *Quarterly Journal of the Royal Meteorological Society*, 141(693):2989–3005, 2015.

[96] T Dunkerton, CP F Hsu, and ME McIntyre. Some eulerian and lagrangian diagnostics for a model stratospheric warming. *Journal of the Atmospheric Sciences*, 38(4):819–844, 1981.

[97] Adam Rupe, Nalini Kumar, Vladislav Epifanov, Karthik Kashinath, Oleksandr Pavlyk, Frank Schlimbach, Mostofa Patwary, Sergey Maidanov, Victor Lee, Mr Prabhat, et al. Disco: Physics-based unsupervised discovery of coherent structures in spatiotemporal systems. In *2019 IEEE/ACM Workshop on Machine Learning in High Performance Computing Environments (MLHPC)*, pages 75–87. IEEE, 2019.

[98] Yunjie Liu, Evan Racah, Joaquin Correa, Amir Khosrowshahi, David Lavers, Kenneth Kunkel, Michael Wehner, William Collins, et al. Application of deep convolutional neural networks for detecting extreme weather in climate datasets. *arXiv preprint arXiv:1605.01156*, 2016.

[99] Evan Racah, Christopher Beckham, Tegan Maharaj, Samira Ebrahimi Kahou, Mr. Prabhat, and Christopher Pal. Extremeweather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. pages 3405–3416, December 2017.

[100] Mayur Mudigonda, Sookyung Kim, Ankur Mahesh, Samira Kahou, Karthik Kashinath, Dean Williams, Vincen Michalski, Travis O'Brien, and Mr Prabhat. Segmenting and tracking extreme climate events using neural networks. In *Deep Learning for Physical Sciences (DLPS) Workshop, held with NIPS Conference*, 2017.

[101] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[102] Swen Brands, José M Gutiérrez, and D San-Martín. Twentieth-century atmospheric river activity along the west coasts of europe and north america: algorithm formulation, reanalysis uncertainty and links to atmospheric circulation patterns. *Climate Dynamics*, 48(9-10):2771–2795, 2017.

[103] N Goldenson, LR Leung, CM Bitz, and E Blanchard-Wrigglesworth. Influence of atmospheric rivers on mountain snowpack in the western united states. *Journal of Climate*, 31(24):9921–9940, 2018.

[104] Bin Guan and Duane E Waliser. Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. *Journal of Geophysical Research: Atmospheres*, 120(24):12514–12535, 2015.

[105] Juan M Lora, Jonathan L Mitchell, Camille Risi, and Aradhna E Tripati. North pacific atmospheric rivers and their influence on western north america at the last glacial maximum. *Geophysical Research Letters*, 44(2):1051–1059, 2017.

[106] Bryan D Mundhenk, Elizabeth A Barnes, and Eric D Maloney. All-season climatology and variability of atmospheric river frequencies over the north pacific. *Journal of Climate*, 29(13):4885–4903, 2016.

[107] Samson Hagos, L Ruby Leung, Qing Yang, Chun Zhao, and Jian Lu. Resolution and dynamical core dependence of atmospheric river frequency in global model simulations. *Journal of Climate*, 28(7):2764–2776, 2015.

[108] David A Lavers, Gabriele Villarini, Richard P Allan, Eric F Wood, and Andrew J Wade. The detection of atmospheric rivers in atmospheric reanalyses and their links to british winter floods and the large-scale climatic circulation. *Journal of Geophysical Research: Atmospheres*, 117(D20), 2012.

[109] Jonathan J Rutz, W James Steenburgh, and F Martin Ralph. Climatological characteristics of atmospheric rivers and their inland penetration over the western united states. *Monthly Weather Review*, 142(2):905–921, 2014.

[110] Christine A Shields and Jeffrey T Kiehl. Simulating the pineapple express in the half degree community climate system model, ccsm4. *Geophysical Research Letters*, 43(14):7767–7773, 2016.

[111] Oliver Rübel, Surendra Byna, Kesheng Wu, Fuyu Li, Michael Wehner, Wes Bethel, et al. Teca: A parallel toolkit for extreme climate analysis. *Procedia Computer Science*, 9:866–876, 2012.

[112] Surendra Prabhat, Byna, Venkatram Vishwanath, Eli Dart, Michael Wehner, William D Collins, et al. Teca: Petascale pattern recognition for climate science. In *International Conference on Computer Analysis of Images and Patterns*, pages 426–436. Springer, 2015.

[113] Yunjie Liu, Evan Racah, Mr Prabhat, Correa Joaquin, Amir Khosrowshahi, David Lavers, Kenneth Kunkel, Michael Wehner, and William Collins. Application of deep convolutional neural networks for detecting extreme weather in climate datasets. In *Advances in Big Data Analytics, The 2016 WorldComp International Conference Proceedings*, 2017.

[114] Scott Lee Sellars, Xiaogang Gao, and Soroosh Sorooshian. An object-oriented approach to investigate impacts of climate oscillations on precipitation: A western united states case study. *Journal of Hydrometeorology*, 16(2):830–842, 2015.

[115] Alexander Gershunov, Tamara Shulgina, F Martin Ralph, David A Lavers, and Jonathan J Rutz. Assessing the climate-scale variability of atmospheric rivers affecting western north america. *Geophysical Research Letters*, 44(15):7900–7908, 2017.

[116] L Ruby Leung and Yun Qian. Atmospheric rivers induced heavy precipitation and flooding in the western us simulated by the wrf regional climate model. *Geophysical research letters*, 36(3), 2009.

[117] Ashley E Payne and Gudrun Magnusdottir. Dynamics of landfalling atmospheric rivers over the north pacific in 30 years of merra reanalysis. *Journal of Climate*, 27(18):7133–7150, 2014.

[118] Ashley E Payne and Gudrun Magnusdottir. An evaluation of atmospheric rivers over the north pacific in cmip5 and their response to warming under rcp 8.5. *Journal of Geophysical Research: Atmospheres*, 120(21):11–173, 2015.

[119] Alexandre M Ramos, Raquel Nieto, Ricardo Tomé, Luis Gimeno, Ricardo M Trigo, Margarida LR Liberato, and David A Lavers. Atmospheric rivers moisture sources from a lagrangian perspective. *Earth System Dynamics*, 7(2):371–384, 2016.

[120] Paul A Ullrich and Colin M Zarzycki. Tempestextremes: a framework for scale-insensitive pointwise feature tracking on unstructured grids. *Geoscientific Model Development*, 10(3):1069, 2017.

[121] Xingying Huang, Daniel L Swain, Daniel B Walton, Samantha Stevenson, and Alex D Hall. Simulating and evaluating atmospheric river-induced precipitation extremes along the us pacific coast: Case studies from 1980–2017. *Journal of Geophysical Research: Atmospheres*, 125(4):e2019JD031554, 2020.

[122] Evan Racah, Christopher Beckham, Tegan Maharaj, Samira Ebrahimi Kahou, Mr Prabhat, and Chris Pal. Extremeweather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. In *Advances in Neural Information Processing Systems*, pages 3402–3413, 2017.

[123] David A Lavers, Ervin Zsoter, David S Richardson, and Florian Pappenberger. An assessment of the ecmwf extreme forecast index for water vapor transport during boreal winter. *Weather and Forecasting*, 32(4):1667–1674, 2017.

[124] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252, 2015.

[125] Simon Jégou, Michal Drozdzal, David Vazquez, Adriana Romero, and Yoshua Bengio. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 11–19, 2017.

[126] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger.

Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[127] SL Sellars, B Kawzenuk, P Nguyen, FM Ralph, and S Sorooshian. Genesis, pathways, and terminations of intense global water vapor transport in association with large-scale climate patterns. *Geophysical Research Letters*, 2017.

[128] EB Garriott. Long range forecasts. *US Weather Bureau Bulletin*, 35, 1904.

[129] R_ Berggren, B Bolin, and C-G Rossby. An aerological study of zonal motion, its perturbations and break-down. *Tellus*, 1(2):14–37, 1949.

[130] Daniel F Rex. The effect of atlantic blocking action upon european climate. *Tellus*, 3(2):100–112, 1951.

[131] Roberto Benzi and AC Wiin-Nielson. Anomalous atmospheric flows and blocking. 1986.

[132] William L Grose and Brian J Hoskins. On the influence of orography on large-scale atmospheric flow. *Journal of the Atmospheric Sciences*, 36(2):223–234, 1979.

[133] Hisashi Nakamura and John M Wallace. Synoptic behavior of baroclinic eddies during the blocking onset. *Monthly weather review*, 121(7):1892–1903, 1993.

[134] Hisashi Nakamura, Mototaka Nakamura, and Jeffrey L Anderson. The role of high-and low-frequency dynamics in blocking formation. *Monthly Weather Review*, 125(9):2074–2093, 1997.

[135] Robert D Elliott and Theodore B Smith. A study of the effects of large blocking highs on the general circulation in the northern-hemisphere westerlies. *Journal of Meteorology*, 6(2):68–85, 1949.

[136] WB White and NE Clark. On the development of blocking ridge activity over the central north pacific. *Journal of the Atmospheric Sciences*, 32(3):489–502, 1975.

[137] JF Austin. The blocking of middle latitude westerly winds by planetary waves. *Quarterly Journal of the Royal Meteorological Society*, 106(448):327–350, 1980.

[138] RA Treidl, EC Birch, and P Sajecki. Blocking action in the northern hemisphere: A climatological study. *Atmosphere-ocean*, 19(1):1–23, 1981.

[139] RM Dole. The objective representation of blocking patterns. *The general circulation: theory, modeling and observations*, pages 406–426, 1978.

[140] JG Charney, J Shukla, and KC Mo. Comparison of a barotropic blocking theory with observation. *Journal of the Atmospheric Sciences*, 38(4):762–779, 1981.

[141] H Van Loon. Blocking action in the southern hemisphere. part i. *Notos*, 5:171–175, 1956.

[142] JJ Taljaard. Synoptic meteorology of the southern hemisphere. In *Meteorology of the Southern Hemisphere*, pages 139–213. Springer, 1972.

[143] Alfred David F Wright. *Blocking action in the Australian region*. Australian Government Publishing Service, 1974.

[144] Harald Lejenäs and Hans Økland. Characteristics of northern hemisphere blocking as determined from a long time series of observational data. *Tellus A*, 35(5):350–362, 1983.

[145] Harald Lejenäs. Characteristics of southern hemisphere blocking as determined from a time series of observational data. *Quarterly Journal of the Royal Meteorological Society*, 110(466):967–979, 1984.

[146] Stefano Tibaldi and Franco Molteni. On the operational predictability of blocking. *Tellus A*, 42(3):343–365, 1990.

[147] S Tibaldi, E Tosi, A Navarra, and L Pedulli. Northern and southern hemisphere seasonal variability of blocking frequency and predictability. *Monthly Weather Review*, 122(9):1971–2003, 1994.

[148] Randall M Dole and Neil D Gordon. Persistent anomalies of the extratropical northern hemisphere wintertime circulation: Geographical distribution and regional persistence characteristics. *Monthly Weather Review*, 111(8):1567–1586, 1983.

[149] Cornelia Schwierz, Mischa Croci-Maspoli, and HC Davies. Perspicacious indicators of atmospheric blocking. *Geophysical research letters*, 31(6), 2004.

[150] Joanna L Pelly and Brian J Hoskins. A new perspective on blocking. *Journal of the atmospheric sciences*, 60(5):743–755, 2003.

[151] Jerome Namias. Characteristics of the general circulation over the northern hemisphere during the abnormal winter 1946–47. *Mon. Wea. Rev*, 75(8):145–152, 1947.

[152] Robert Ghrist. Barcodes: the persistent topology of data. *Bulletin of the American Mathematical Society*, 45(1):61–75, 2008.

[153] Gunnar Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

[154] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[155] John E. Hopcroft and Jeffrey D. Ullman. Set merging algorithms. *SIAM Journal on Computing*, 2(4):294–303, 1973.

[156] Robert Endre Tarjan. Efficiency of a good but not linear set union algorithm. *Journal of the ACM (JACM)*, 22(2):215–225, 1975.

[157] Herbert Edelsbrunner and John Harer. *Computational topology: an introduction*. American Mathematical Soc., 2010.

[158] Chih-Chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3):27, 2011.

[159] Herbert Edelsbrunner and Dmitriy Morozov. Persistent homology: Theory and practice. Technical report, Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA (US), 2012.

[160] Christopher JC Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167, 1998.

[161] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. A practical guide to support vector classification. 2003.

[162] Haibo He and Edwardo A Garcia. Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, 21(9):1263–1284, 2009.

[163] Guillaume Lemaître, Fernando Nogueira, and Christos K. Aridas. Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. *Journal of Machine Learning Research*, 18(17):1–5, 2017.

[164] Brian Eaton. User's guide to the community atmosphere model cam-5.1, 2011.

[165] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *Proceedings of the International Conference on Learning Representation (ICLR*, 2015.

[166] Emo Welzl. Smallest enclosing disks (balls and ellipsoids). In *New results and new trends in computer science*, pages 359–370. Springer, 1991.

[167] Yann LeCun, Bernhard E Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne E Hubbard, and Lawrence D Jackel. Handwritten digit recognition with a back-propagation network. In *Advances in neural information processing systems*, pages 396–404, 1990.

[168] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

[169] Y-Lan Boureau, Jean Ponce, and Yann LeCun. A theoretical analysis of feature pooling in visual recognition. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 111–118, 2010.

[170] Markus Reichstein, Gustau Camps-Valls, Bjorn Stevens, Martin Jung, Joachim Denzler, Nuno Carvalhais, et al. Deep learning and process understanding for data-driven earth system science. *Nature*, 566(7743):195–204, 2019.

[171] Léon Bottou. Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010*, pages 177–186. Springer, 2010.

[172] Quinn McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157, 1947.

[173] Frank Wilcoxon. Probability tables for individual comparisons by ranking methods. *Biometrics*, 3(3):119–122, 1947.

[174] Jason M Wiedenmann, Anthony R Lupo, Igor I Mokhov, and Elena A Tikhonova. The climatology of blocking anticyclones for the northern and southern hemispheres: Block intensity as a diagnostic. *Journal of Climate*, 15(23):3459–3473, 2002.

[175] Amir Shabbar, Jianping Huang, and Kaz Higuchi. The relationship between the wintertime north atlantic oscillation and blocking episodes in the north atlantic. *International Journal of Climatology*, 21(3):355–369, 2001.

[176] Mischa Croci-Maspoli, Cornelia Schwierz, and Huw C Davies. Atmospheric blocking: Space-time links to the nao and pna. *Climate Dynamics*, 29(7-8):713–725, 2007.

[177] Karthik Prabhat, Kashinath, Mayur Mudigonda, Sol Kim, Lukas Kapp-Schwoerer, Andre Graubner, Ege Karaismailoglu, Leo von Kleist, Thorsten Kurth, Annette Greiner, Kevin Yang, et al. Climatenet: an expert-labelled open dataset and deep learning architecture for enabling high-precision analyses of extreme weather. *Geoscientific Model Development Discussions*, pages 1–28, 2020.